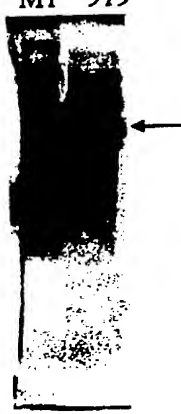


PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/31, C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N 33/50</b>		<b>A2</b>	(11) International Publication Number: <b>WO 99/57280</b> (43) International Publication Date: 11 November 1999 (11.11.99)
(21) International Application Number: PCT/US99/09346 (22) International Filing Date: 30 April 1999 (30.04.99) (30) Priority Data: 60/083,758 1 May 1998 (01.05.98) US 60/094,869 31 July 1998 (31.07.98) US 60/098,994 2 September 1998 (02.09.98) US 60/099,062 2 September 1998 (02.09.98) US 60/103,749 9 October 1998 (09.10.98) US 60/103,794 9 October 1998 (09.10.98) US 60/103,796 9 October 1998 (09.10.98) US 60/121,528 25 February 1999 (25.02.99) US (71) Applicants (for all designated States except US): CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE INSTITUTE FOR GENOMIC RESEARCH [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): FRASER, Claire [US/US]; Rockville, MD (US). GALEOTTI, Cesira [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). GRANDI, Guido [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). HICKEY, Erin [US/US]; Gaithersburg, MD (US).		(US). MASIGNANI, Vega [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). MORA, Marirosa [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). PETERSEN, Jeremy [US/US]; Arlington, VA (US). PIZZA, Mariagrazia [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RAPPUOLI, Rino [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RATTI, Giulio [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCALATO, Enzo [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCARSELLI, Maria [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). TETTELIN, Herve [US/US]; Gaithersburg, MD (US). VENTER, J., Craig [US/US]; Rockville, MD (US). (74) Agent: HARBIN, Alisa, A.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> Without international search report and to be republished upon receipt of that report.	
(54) Title: NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS			
(57) Abstract The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.			
<div style="text-align: right;"><b>919 (46 kDa)</b></div> <div style="text-align: right;"><b>A) PURIFICATION</b></div> <div style="text-align: right;"><b>M1 919</b> </div>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TC	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			P*	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		



30/PR+5

09/674546  
526 Rec'd PCT/PTO 01 NOV 2000  
PCT/US99/09346

## NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

This application is a continuation-in-part of the following U.S. Provisional Patent applications, from each of which priority is claimed, and each of which is incorporated by reference in its entirety: 60/083,758 (filed May 1, 1998); 60/094,869 (filed July 31, 1998); 60/098,994 (filed September 2, 1998); 60/099,062 (filed September 2, 1998); 60/103,749 (filed October 9, 1998); 60/103,794 (filed October 9, 1998); 60/103,796 (filed October 9, 1998); and 60/121,528 (filed February 25, 1999).

This invention relates to antigens from the bacterial species: *Neisseria meningitidis* and *Neisseria gonorrhoeae*.

### BACKGROUND

*Neisseria meningitidis* is a non-motile, gram negative diplococcus human pathogen. It colonizes the pharynx, causing meningitis and, occasionally, septicaemia in the absence of meningitis. It is closely related to *N. gonorrhoea*. although one feature that clearly differentiates meningococcus from gonococcus is the presence of a polysaccharide capsule that is present in all pathogenic meningococci.

*N. meningitidis* causes both endemic and epidemic disease. In the United States the attack rate is 0.6-1 per 100,000 persons per year, and it can be much greater during outbreaks. (see Lieberman *et al.* (1996) Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. *JAMA* 275(19):1499-1503; Schuchat *et al* (1997) Bacterial Meningitis in the United States in 1995. *N Engl J Med* 337(14):970-976). In developing countries, endemic disease rates are much higher and during epidemics incidence rates can reach 500 cases per 100,000 persons per year. Mortality is extremely high, at 10-20% in the United States, and much higher in developing countries. Following the introduction of the conjugate vaccine against *Haemophilus influenzae*, *N. meningitidis* is the major cause of bacterial meningitis at all ages in the United States (Schuchat *et al* (1997) *supra*).

Based on the organism's capsular polysaccharide, 12 serogroups of *N. meningitidis* have been identified. Group A is the pathogen most often implicated in epidemic disease in sub-Saharan Africa. Serogroups B and C are responsible for the vast majority of cases in the

United States and in most developed countries. Serogroups W135 and Y are responsible for the rest of the cases in the United States and developed countries. The meningococcal vaccine currently in use is a tetravalent polysaccharide vaccine composed of serogroups A, C, Y and W135. Although efficacious in adolescents and adults, it induces a poor immune response and short duration of protection, and cannot be used in infants [eg. Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. This is because polysaccharides are T-cell independent antigens that induce a weak immune response that cannot be boosted by repeated immunization. Following the success of the vaccination against *H.influenzae*, conjugate vaccines against serogroups A and C have been developed and are at the final stage of clinical testing (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease". In: *New Generation Vaccines, supra*, pp. 469-488; Lieberman *et al* (1996) *supra*; Costantino *et al* (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. *Vaccine* 10:691-698).

Meningococcus B (menB) remains a problem, however. This serotype currently is responsible for approximately 50% of total meningitis in the United States, Europe, and South America. The polysaccharide approach cannot be used because the menB capsular polysaccharide is a polymer of  $\alpha(2-8)$ -linked *N*-acetyl neuraminic acid that is also present in mammalian tissue. This results in tolerance to the antigen; indeed, if an immune response were elicited, it would be anti-self, and therefore undesirable. In order to avoid induction of autoimmunity and to induce a protective immune response, the capsular polysaccharide has, for instance, been chemically modified substituting the *N*-acetyl groups with *N*-propionyl groups, leaving the specific antigenicity unaltered (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? *Clin Microbiol Rev* 7(4):559-575).

Alternative approaches to menB vaccines have used complex mixtures of outer membrane proteins (OMPs), containing either the OMPs alone, or OMPs enriched in porins, or deleted of the class 4 OMPs that are believed to induce antibodies that block bactericidal activity. This approach produces vaccines that are not well characterized. They are able to protect against the homologous strain, but are not effective at large where there are many antigenic variants of the outer membrane proteins. To overcome the antigenic variability, multivalent vaccines containing up to nine different porins have been constructed (eg. Poolman JT (1992) Development of a meningococcal vaccine. *Infect. Agents Dis.* 4:13-28).

Additional proteins to be used in outer membrane vaccines have been the opa and opc proteins, but none of these approaches have been able to overcome the antigenic variability (eg. Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. *Vaccine* 14(1):49-53).

A certain amount of sequence data is available for meningococcal and gonococcal genes and proteins (e.g. EP-A-0467714, WO96/29412), but this is by no means complete. Other men B proteins may include those listed in WO 97/28273, WO 96/29412, WO 95/03413, US 5,439,808, and US 5,879,686.

The provision of further sequences could provide an opportunity to identify secreted or surface-exposed proteins that are presumed targets for the immune system and which are not antigenically variable. For instance, some of the identified proteins could be components of efficacious vaccines against meningococcus B, some could be components of vaccines against all meningococcal serotypes, and others could be components of vaccines against all pathogenic *Neisseriae* including *Neisseria meningitidis* or *Neisseria gonorrhoeae*. Those sequences specific to *N. meningitidis* or *N. gonorrhoeae* that are more highly conserved are further preferred sequences.

It is thus an object of the invention is to provide Neisserial DNA sequences which encode proteins that are antigenic or immunogenic.

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the products of protein expression and purification of the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 2 illustrates the products of protein expression and purification of the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 3 illustrates the products of protein expression and purification of the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 4 illustrates the products of protein expression and purification of the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 5 illustrates the products of protein expression and purification of the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 6 illustrates the products of protein expression and purification of the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 7 illustrates the products of protein expression and purification of the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 8 illustrates the products of protein expression and purification of the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 9 illustrates the products of protein expression and purification of the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 10 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 11 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 12 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 13 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 14 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 15 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 16 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 17 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 18 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 19 shows an alignment comparison of amino acid sequences for ORF 225 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 20 shows an alignment comparison of amino acid sequences for ORF 235 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 21 shows an alignment comparison of amino acid sequences for ORF 287 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 22 shows an alignment comparison of amino acid sequences for ORF 519 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 23 shows an alignment comparison of amino acid sequences for ORF 919 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

## THE INVENTION

The invention provides proteins comprising the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples.

It also provides proteins comprising sequences homologous (i.e., those having sequence identity) to the *N. meningitidis* amino acid sequences disclosed in the examples. Depending on the particular sequence, the degree of homology (sequence identity) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). These proteins include mutants and allelic variants of the sequences disclosed in the examples. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman

homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with parameters: gap penalty 12, gap extension penalty 1.

The invention further provides proteins comprising fragments of the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples. The fragments should comprise at least *n* consecutive amino acids from the sequences and, depending on the particular sequence, *n* is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20 or more). Preferably the fragments comprise an epitope from the sequence.

The proteins of the invention can, of course, be prepared by various means (eg. recombinant expression, purification from cell culture, chemical synthesis *etc.*) and in various forms (eg. native, fusions *etc.*). They are preferably prepared in substantially pure or isolated form (*ie.* substantially free from other *N. meningitidis* or *N. gonorrhoeae* host cell proteins).

According to a further aspect, the invention provides antibodies which bind to these proteins. These may be polyclonal or monoclonal and may be produced by any suitable means.

According to a further aspect, the invention provides nucleic acid comprising the *N. meningitidis* nucleotide sequences and *N. gonorrhoeae* nucleotide sequences disclosed in the examples.

According to a further aspect, the invention comprises nucleic acids having sequence identity of greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) to the nucleic acid sequences herein. Sequence identity is determined as above-discussed.

According to a further aspect, the invention comprises nucleic acid that hybridizes to the sequences provided herein. Conditions for hybridization are set forth herein.

Nucleic acid comprising fragments of these sequences are also provided. These should comprise at least *n* consecutive nucleotides from the *N. meningitidis* sequences or *N. gonorrhoeae* sequences and depending on the particular sequence, *n* is 10 or more (eg 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

According to a further aspect, the invention provides nucleic acid encoding the proteins and protein fragments of the invention.

It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (eg. for antisense or probing purposes).

Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, in part or in whole, from genomic or cDNA libraries, from the

organism itself *etc.*) and can take various forms (*eg.* single stranded, double stranded, vectors, probes *etc.*).

In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also protein nucleic acids (PNA) *etc.*

According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (*eg.* expression vectors) and host cells transformed with such vectors.

According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as diagnostic reagents or as immunogenic compositions.

The invention also provides nucleic acid, protein, or antibody according to the invention for use as medicaments (*eg.* as vaccines) or as diagnostic reagents. It also provides the use of nucleic acid, protein, or antibody according to the invention in the manufacture of (I) a medicament for treating or preventing infection due to Neisserial bacteria (ii) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria or (iii) for raising antibodies. Said Neisserial bacteria may be any species or strain (such as *N. gonorrhoeae*) but are preferably *N. meningitidis*, especially strain B or strain C.

The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention.

According to further aspects, the invention provides various processes.

A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression.

A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes.

A process for detecting proteins of the invention is provided, comprising the steps of: (a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

#### Methodology - Summary of standard procedures and techniques.

##### General

This invention provides *Neisseria meningitidis* menB nucleotide sequences, amino acid sequences encoded therein. With these disclosed sequences, nucleic acid probe assays and expression cassettes and vectors can be produced. The expression vectors can be transformed into host cells to produce proteins. The purified or isolated polypeptides (which may also be chemically synthesized) can be used to produce antibodies to detect menB proteins. Also, the host cells or extracts can be utilized for biological assays to isolate agonists or antagonists. In addition, with these sequences one can search to identify open reading frames and identify amino acid sequences. The proteins may also be used in immunogenic compositions, antigenic compositions and as vaccine components.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature e.g., Sambrook *Molecular Cloning; A Laboratory Manual, Second Edition* (1989); *DNA Cloning, Volumes I and ii* (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); *Transcription and Translation* (B.D. Hames & S.J. Higgins eds. 1984); *Animal Cell Culture* (R.I. Freshney ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the *Methods in Enzymology* series (Academic Press, Inc.), especially volumes 154 & 155; *Gene Transfer Vectors for Mammalian Cells* (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer-Verlag, N.Y.), and *Handbook of Experimental Immunology, Volumes I-IV* (D.M. Weir and C.C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification.



All publications, patents, and patent applications cited herein are incorporated in full by reference.

#### Expression systems

The *Neisseria* menB nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, plant cells, baculoviruses, bacteria, and yeast.

##### i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells." In *Molecular Cloning: A Laboratory Manual*, 2nd ed.).

Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible). Depending on the promoter selected, many promoters may be inducible using known substrates, such as the use of the mouse mammary tumor virus (MMTV) promoter with the glucocorticoid responsive element (GRE) that is induced by glucocorticoid in hormone-responsive transformed cells (see for example, U.S. Patent 5,783,681).

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a

regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) *Science* 236:1237; Alberts et al. (1989) *Molecular Biology of the Cell*, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) *EMBO J.* 4:761) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) *Proc. Natl. Acad. Sci.* 79:6777) and from human cytomegalovirus (Boshart et al. (1985) *Cell* 41:521). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) *Trends Genet.* 2:215; Maniatis et al. (1987) *Science* 236:1237).

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birstiel et al. (1985) *Cell* 41:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In *Transcription and splicing* (ed. B.D. Hames and D.M.

Glover); Proudfoot (1989) *Trends Biochem. Sci.* 14:105). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In *Molecular Cloning: A Laboratory Manual*).

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 (Gluzman (1981) *Cell* 23:175) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) *Mol. Cell. Biol.* 9:946) and pHEBO (Shimizu et al. (1986) *Mol. Cell. Biol.* 6:1074).

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

## ii. Plant Cellular Expression Systems

There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: U.S. 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, *Phytochemistry* 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., *Nucleic Acids Research* 15:2515-2535 (1987); Wirsal et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: *Advanced Plant Physiology*, Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel and Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987)

Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for *Agrobacterium* transformations, T DNA sequences for *Agrobacterium*-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, *Plant Mol. Biol. Repr.* 11(2):165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

Since the ultimate expression of the desired gene product will be in a eucaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code, Reed and Maniatis, *Cell* 41:95-105, 1985.

The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA. Crossway, *Mol. Gen. Genet.*, 202:179-185, 1985. The genetic material may also be transferred into the plant cell by using polyethylene glycol, Krens, et al., *Nature*, 296, 72-74, 1982. Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface, Klein, et al., *Nature*, 327, 70-73, 1987 and Knudsen and Muller, 1991, *Planta*, 185:330-336 teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

The vector may also be introduced into the plant cells by electroporation. (Fromm et al., *Proc. Natl. Acad. Sci. USA* 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and *Datura*.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension.

These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected. Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

### iii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to

those skilled in the art and fully described in Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its own set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, *Virology* (1989) 17:31.

The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) *Ann. Rev. Microbiol.*, 42:177) and a prokaryotic ampicillin-resistance (*amp*) gene and origin of replication for selection and propagation in *E. coli*.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: *The Molecular Biology of Baculoviruses* (ed. Walter Doerfler); EPO



Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak et al., (1988), *J. Gen. Virol.* 69:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) *Gene*, 73:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human (alpha)  $\alpha$ -interferon, Maeda et al., (1985), *Nature* 315:592; human gastrin-releasing peptide, Lebacqz-Verheyden et al., (1988), *Molec. Cell. Biol.* 8:3129; human IL-2, Smith et al., (1985) *Proc. Nat'l Acad. Sci. USA*, 82:8404; mouse IL-3, (Miyajima et al., (1987) *Gene* 58:273; and human glucocerebrosidase, Martin et al. (1988) *DNA*, 7:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by *in vitro* incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See

Summers and Smith *supra*; Ju et al. (1987); Smith et al., *Mol. Cell. Biol.* (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), *Bioessays* 4:91. The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15  $\mu\text{m}$  in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. *Current Protocols in Microbiology* Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, *supra*; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, *inter alia*: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) *J. Virol.* 56:153; Wright (1986) *Nature* 321:718; Smith et al., (1983) *Mol. Cell. Biol.* 3:2156; and see generally, Fraser, et al. (1989) *In Vitro Cell. Dev. Biol.* 25:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith *supra*.

The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

#### iv. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps

initiate transcription of the lac operon in *Escherichia coli* (*E. coli*) (Raibaud *et al.* (1984) *Annu. Rev. Genet.* 18:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) (Chang *et al.* (1977) *Nature* 198:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (*trp*) (Goeddel *et al.* (1980) *Nuc. Acids Res.* 8:4057; Yelverton *et al.* (1981) *Nucl. Acids Res.* 9:731; U.S. Patent 4,738,921; EPO Publ. Nos. 036 776 and 121 775). The beta-lactamase (*bla*) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In *Interferon 3* (ed. I. Gresser)), bacteriophage lambda PL (Shimatake *et al.* (1981) *Nature* 292:128) and T5 (U.S. Patent 4,689,406) promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Patent 4,551,433). For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor (Amann *et al.* (1983) *Gene* 25:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci.* 80:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier *et al.* (1986) *J. Mol. Biol.* 189:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci.* 82:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon

(Shine *et al.* (1975) *Nature* 254:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of *E. coli* 16S rRNA (Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberg)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site, it is often necessary to optimize the distance between the SD sequence and the ATG of the eukaryotic gene (Sambrook *et al.* (1989) "Expression of cloned genes in *Escherichia coli*." In *Molecular Cloning: A Laboratory Manual*).

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* or *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai *et al.* (1984) *Nature* 309:810). Fusion proteins can also be made with sequences from the *lacZ* (Jia *et al.* (1987) *Gene* 60:197), *trpE* (Allen *et al.* (1987) *J. Biotechnol.* 5:93; Makoff *et al.* (1989) *J. Gen. Microbiol.* 135:11), and *Chey* (EPO Publ. No. 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller *et al.* (1989) *Bio/Technology* 7:698).

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (U.S. Patent 4,336,336). The

signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E. coli* outer membrane protein gene (*ompA*) (Masui *et al.* (1983), in: *Experimental Manipulation of Gene Expression*; Ghrayeb *et al.* (1984) *EMBO J.* 3:2437) and the *E. coli* alkaline phosphatase signal sequence (*phoA*) (Oka *et al.* (1985) *Proc. Natl. Acad. Sci.* 82:7212). As an additional example, the signal sequence of the alpha-amylase gene from various *Bacillus* strains can be used to secrete heterologous proteins from *B. subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. No. 244 042).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number

vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various *Bacillus* strains integrate into the *Bacillus* chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies *et al.* (1978) *Annu. Rev. Microbiol.* 32:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: *Bacillus subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541), *Escherichia coli* (Shimatake *et al.* (1981) *Nature* 292:128; Amann *et al.* (1985) *Gene* 40:183; Studier *et al.* (1986) *J. Mol. Biol.* 189:113; EPO Publ. Nos. 036 776, 136 829 and 136 907), *Streptococcus cremoris* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655); *Streptococcus lividans* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655), *Streptomyces lividans* (U.S. Patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with CaCl<sub>2</sub> or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by

electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See e.g., use of *Bacillus*: Masson *et al.* (1989) *FEMS Microbiol. Lett.* 60:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541; use of *Campylobacter*: Miller *et al.* (1988) *Proc. Natl. Acad. Sci.* 85:856; and Wang *et al.* (1990) *J. Bacteriol.* 172:949; use of *Escherichia coli*: Cohen *et al.* (1973) *Proc. Natl. Acad. Sci.* 69:2110; Dower *et al.* (1988) *Nucleic Acids Res.* 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. In *Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol.* 53:159; Taketo (1988) *Biochim. Biophys. Acta* 949:318; use of *Lactobacillus*: Chassy *et al.* (1987) *FEMS Microbiol. Lett.* 44:173; use of *Pseudomonas*: Fiedler *et al.* (1988) *Anal. Biochem* 170:38; use of *Staphylococcus*: Augustin *et al.* (1990) *FEMS Microbiol. Lett.* 66:203; use of *Streptococcus*: Barany *et al.* (1980) *J. Bacteriol.* 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun.* 32:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655; Somkuti *et al.* (1987) *Proc. 4th Eur. Cong. Biotechnology* 1:412.

#### v. Yeast Expression

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences.



Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the *ADH2*, *GAL4*, *GAL10*, OR *PHO5* genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, *inter alia*, (Cohen *et al.* (1980) *Proc. Natl. Acad. Sci. USA* 77:1078; Henikoff *et al.* (1981) *Nature* 283:835; Hollenberg *et al.* (1981) *Curr. Topics Microbiol. Immunol.* 96:119; Hollenberg *et al.* (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast *Saccharomyces cerevisiae*," in: *Plasmids of Medical, Environmental and Commercial Importance* (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon *et al.* (1980) *Gene* 11:163; Panthier *et al.* (1980) *Curr. Genet.* 2:109;).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, plant, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human

superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62:096,086) and the A-factor gene (U.S. Patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator

sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.* (1979) *Gene* 8:17-24), pCl/1 (Brake *et al.* (1984) *Proc. Natl. Acad. Sci. USA* 81:4642-4646), and YRp17 (Stinchcomb *et al.* (1982) *J. Mol. Biol.* 158:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Enter a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake *et al.*, *supra*.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver *et al.* (1983) *Methods in Enzymol.* 101:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver *et al.*, *supra*. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as *ADE2*, *HIS4*, *LEU2*, *TRP1*, and *ALG7*, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of *CUP1* allows yeast to grow in the presence of copper ions (Butt *et al.* (1987) *Microbiol. Rev.* 51:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors and methods of introducing exogenous DNA into yeast hosts have been developed for, *inter alia*, the following yeasts: *Candida albicans* (Kurtz, *et al.* (1986) *Mol. Cell. Biol.* 6:142); *Candida maltosa* (Kunze, *et al.* (1985) *J. Basic Microbiol.* 25:141); *Hansenula polymorpha* (Gleeson, *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302); *Kluyveromyces fragilis* (Das, *et al.* (1984) *J. Bacteriol.* 158:1165); *Kluyveromyces lactis* (De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:737; Van den Berg *et al.* (1990) *Bio/Technology* 8:135); *Pichia guilliermondii* (Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141); *Pichia pastoris* (Cregg, *et al.* (1985) *Mol. Cell. Biol.* 5:3376; U.S. Patent Nos. 4,837,148 and 4,929,555); *Saccharomyces cerevisiae* (Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163); *Schizosaccharomyces pombe* (Beach and Nurse (1981) *Nature* 300:706); and *Yarrowia lipolytica* (Davidow, *et al.* (1985) *Curr. Genet.* 10:380471 Gaillardin, *et al.* (1985) *Curr. Genet.* 10:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz *et al.* (1986) *Mol. Cell. Biol.* 6:142; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; *Candida*]; [Gleeson *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302; *Hansenula*]; [Das *et al.* (1984) *J.*

*Bacteriol.* 158:1165; De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:1165; Van den Berg *et al.* (1990) *Bio/Technology* 8:135; Kluyveromyces]; [Cregg *et al.* (1985) *Mol. Cell. Biol.* 5:3376; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; *Pichia*]; [Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163 *Saccharomyces*]; [Beach and Nurse (1981) *Nature* 300:706; *Schizosaccharomyces*]; [Davidow *et al.* (1985) *Curr. Genet.* 10:39; Gaillardin *et al.* (1985) *Curr. Genet.* 10:49; *Yarrowia*].

### Definitions

A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

A "conserved" *Neisseria* amino acid fragment or protein is one that is present in a particular *Neisseria* protein in at least x% of *Neisseria*. The value of x may be 50% or more, e.g., 66%, 75%, 80%, 90%, 95% or even 100% (i.e. the amino acid is found in the protein in question in all *Neisseria*). In order to determine whether an amino acid is "conserved" in a particular *Neisseria* protein, it is necessary to compare that amino acid residue in the sequences of the protein in question from a plurality of different *Neisseria* (a reference population). The reference population may include a number of different *Neisseria* species or may include a single species. The reference population may include a number of different serogroups of a particular species or a single serogroup. A preferred reference population consists of the 5 most common *Neisseria* strains.

The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a *Neisseria* sequence is heterologous to a mouse host cell.

"Epitope" means antigenic determinant, and may elicit a cellular and/or humoral response.

Conditions for "high stringency" are 65 degrees C in 0.1 xSSC 0.5% SDS solution.

An "origin of replication" is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

A "mutant" sequence is defined as a DNA, RNA or amino acid sequence differing from but having homology with the native or disclosed sequence. Depending on the particular sequence, the degree of homology (sequence identity) between the native or disclosed sequence and the mutant sequence is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) which is calculated as described above. As used herein, an "allelic variant" of a nucleic acid molecule, or region, for which nucleic acid sequence is provided herein is a nucleic acid molecule, or region, that occurs at essentially the same locus in the genome of another or second isolate, and that, due to natural variation caused by, for example, mutation or recombination, has a similar but not identical nucleic acid sequence. A coding region allelic variant typically encodes a protein having similar activity to that of the protein encoded by the gene to which it is being compared. An allelic variant can also comprise an alteration in the 5' or 3' untranslated regions of the gene, such as in regulatory control regions. (see, for example, U.S. Patent 5,753,235).

#### Antibodies

As used herein, the term "antibody" refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An "antibody combining site" is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. "Antibody" includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying *Neisseria menB* proteins.

Antibodies elicited against the proteins of the present invention bind to antigenic polypeptides or proteins or protein fragments that are present and specifically associated with strains of *Neisseria meningitidis* menB. In some instances, these antigens may be associated with specific strains, such as those antigens specific for the menB strains. The antibodies of the invention may be immobilized to a matrix and utilized in an immunoassay or on an affinity chromatography column, to enable the detection and/or separation of polypeptides, proteins or protein fragments or cells comprising such polypeptides, proteins or protein fragments. Alternatively, such polypeptides, proteins or protein fragments may be immobilized so as to detect antibodies bindably specific thereto.

Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000g for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (*Nature* (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and

are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly  $^{32}\text{P}$  and  $^{125}\text{I}$ ), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example,  $^{125}\text{I}$  may serve as a radioactive label or as an electron-dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a MAb with biotin, and detect its presence with avidin labeled with  $^{125}\text{I}$ , or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

Antigens, immunogens, polypeptides, proteins or protein fragments of the present invention elicit formation of specific binding partner antibodies. These antigens, immunogens, polypeptides, proteins or protein fragments of the present invention comprise immunogenic compositions of the present invention. Such immunogenic compositions may further comprise or include adjuvants, carriers, or other compositions that promote or enhance



or stabilize the antigens, polypeptides, proteins or protein fragments of the present invention. Such adjuvants and carriers will be readily apparent to those of ordinary skill in the art.

#### Pharmaceutical Compositions

Pharmaceutical compositions can comprise (include) either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature, when given to a patient that is febrile. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

For purposes of the present invention, an effective dose will be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of

organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

#### Delivery Methods

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal and transcutaneous applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

#### Vaccines

Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

Such vaccines comprise immunizing antigen(s) or immunogen(s), immunogenic polypeptide, protein(s) or protein fragments, or nucleic acids (e.g., ribonucleic acid or deoxyribonucleic acid), usually in combination with "pharmaceutically acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or

liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the immunogen or antigen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, etc. pathogens.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalene, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) Ribi<sup>TM</sup> adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox<sup>TM</sup>); (3) saponin adjuvants, such as Stimulon<sup>TM</sup> (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, *etc.*), interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Alum and MF59 are preferred.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), *etc.*

The vaccine compositions comprising immunogenic compositions (e.g., which may include the antigen, pharmaceutically acceptable carrier, and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Alternatively, vaccine compositions comprising immunogenic compositions may comprise an antigen, polypeptide, protein, protein fragment or nucleic acid in a pharmaceutically acceptable carrier.

More specifically, vaccines comprising immunogenic compositions comprise an immunologically effective amount of the immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, *etc.*), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Typically, the vaccine compositions or immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above under pharmaceutically acceptable carriers.

The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal and transcutaneous applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648).

### Gene Delivery Vehicles

Gene therapy vehicles for delivery of constructs, including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches in *in vivo* or *ex vivo* modality. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence *in vivo* can be either constitutive or regulated.

The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; and Kaplitt (1994) *Nature Genetics* 6:148-153.

Retroviral vectors are well known in the art, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) *J. Virol.* 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) *J. Virol.* 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985.

Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53 (1993) 83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci* 81:6349; and Miller (1990) *Human Gene Therapy* 1.

Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) *Biotechniques* 6:616 and Rosenfeld (1991) *Science* 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102,

WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curiel (1992) *Hum. Gene Ther.* 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN-1, both of which are disclosed in Nahreini (1993) *Gene* 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) *J. Virol.* 61:3096). Another exemplary AAV vector is the Double-D ITR vector. Construction of the Double-D ITR vector is disclosed in US Patent 5,478,745. Still other vectors are those disclosed in Carter US Patent 4,797,368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) *Human Gene Therapy* 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

The gene therapy vectors comprising sequences of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVlac described in Geller (1988) *Science* 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ

described in Fink (1992) *Human Gene Therapy* 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR-1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), and those described in US patents 5,091,309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in U.S. Serial No. 08/405,627, filed March 15, 1995, WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

DNA vector systems such as eukaryotic layered expression systems are also useful for expressing the nucleic acids of the invention. See WO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, *Nature* 339 (1989) 385 and Sabin (1973) *J. Biol. Standardization* 1:115; rhinovirus, for example ATCC VR-1110 and those described in Arnold (1990) *J Cell Biochem* L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-111 and ATCC VR-2010 and those described in Fisher-Hoch (1989) *Proc Natl Acad Sci* 86:317; Flexner (1989) *Ann NY Acad Sci* 569:86, Flexner (1990) *Vaccine* 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) *Nature* 277:108 and Madzak (1992) *J Gen Virol* 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) *Proc Natl Acad Sci* 87:3802-3805; Enami & Palese (1991) *J Virol* 65:2711-2713 and Luytjes (1989) *Cell* 59:110, (see also McMichael (1983) *NEJ Med* 309:13, and Yap (1978) *Nature* 273:238 and *Nature* (1979) 277:108); human



immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) *J. Virol.* 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR-600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1244; Ndumu virus, for example ATCC VR-371; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Trinita virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-65 and ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) *Proc Soc Exp Biol Med* 121:190.

Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 08/366,787, filed December 30, 1994 and Curiel (1992) *Hum Gene Ther* 3:147-154 ligand linked DNA, for example see Wu (1989) *J Biol Chem* 264:16985-16987, eucaryotic cell delivery vehicles cells, for example see US Serial No.08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/11033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) *Mol Cell Biol* 14:2411-2418 and in Woffendin (1994) *Proc Natl Acad Sci* 91:1581-1585.

Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu &

Wu (1987) *J. Biol. Chem.* 262:4429-4432, insulin as described in Hucked (1990) *Biochem Pharmacol* 40:253-263, galactose as described in Plank (1992) *Bioconjugate Chem* 3:533-539, lactose or transferrin.

Naked DNA may also be employed to transform a host cell. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

Liposomes that can act as gene delivery vehicles are described in U.S. 5,422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in USSN. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin. Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin *et al* (1994) *Proc. Natl. Acad. Sci. USA* 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. 5,206,152 and WO92/11033.

Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; in WO 95/13796; WO94/23697; and WO91/14445; in EP-0524968; and in Stryer, *Biochemistry*, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) *Biochem Biophys Acta* 600:1; Bayer (1979) *Biochem Biophys Acta* 550:464; Rivnay (1987) *Meth Enzymol* 149:119; Wang (1987) *Proc Natl Acad Sci* 84:7851; Plant (1989) *Anal Biochem* 176:420.

A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0.01 mg/ kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

### **Delivery Methods**

Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered *ex vivo*, to cells derived from the subject; or (3) *in vitro* for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a tumor or lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in *eg.* WO93/14778. Examples of cells useful in *ex vivo* applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells.

Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

### **Polynucleotide and polypeptide pharmaceutical compositions**

In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

#### A. Polypeptides

One example are polypeptides which include, without limitation: asialoglycoprotein (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF),

granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII.

#### B. Hormones, Vitamins, Etc.

Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

#### C. Polyalkylenes, Polysaccharides, etc.

Also, polyalkylene glycol can be included with the desired polynucleotides or polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethylene glycol. In addition, mono-, di-, or polysaccharides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

#### D. Lipids, and Liposomes

The desired polynucleotide or polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide or polypeptide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); and purified transcription factors (Debs (1990) *J. Biol. Chem.* 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner *supra*). Other commercially available liposomes include transfectace (DDAB/DOPE) and

DOTAP/DOPE (Boehringer). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, eg. Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See eg. Straubinger (1983) *Meth. Immunol.* 101:512-527; Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; Papahadjopoulos (1975) *Biochim. Biophys. Acta* 394:483; Wilson (1979) *Cell* 17:77; Deamer & Bangham (1976) *Biochim. Biophys. Acta* 443:629; Ostro (1977) *Biochem. Biophys. Res. Commun.* 76:836; Fraley (1979) *Proc. Natl. Acad. Sci. USA* 76:3348; Enoch & Strittmatter (1979) *Proc. Natl. Acad. Sci. USA* 76:145; Fraley (1980) *J. Biol. Chem.* (1980) 255:10431; Szoka & Papahadjopoulos (1978) *Proc. Natl. Acad. Sci. USA* 75:145; and Schaefer-Ridder (1982) *Science* 215:166.

#### E. Lipoproteins

In addition, lipoproteins can be included with the polynucleotide or polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E.

The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) *Annu Rev. Biochem* 54:699; Law (1986) *Adv. Exp Med. Biol.* 151:162; Chen (1986) *J Biol Chem* 261:12918; Kane (1980) *Proc Natl Acad Sci USA* 77:2465; and Utermann (1984) *Hum Genet* 65:232.

Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phospholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in *Meth. Enzymol.* 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in *Meth. Enzymol. (supra)*; Pitas (1980) *J. Biochem.* 255:5454-5460 and Mahey (1979) *J Clin. Invest* 64:743-750.

Lipoproteins can also be produced by *in vitro* or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) *Annu Rev Biophys Chem* 15:403 and Radding (1958) *Biochim Biophys Acta* 30: 443.

Lipoproteins can also be purchased from commercial suppliers, such as Biomedical Technologies, Inc., Stoughton, Massachusetts, USA.

Further description of lipoproteins can be found in Zuckermann et al., PCT. Appln. No. US97/14465.

#### F. Polycationic Agents

Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide or polypeptide to be delivered.

Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both in vitro, ex vivo, and in vivo applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc.

The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non-histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic acid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

Organic polycationic agents include: spermine, spermidine, and putrescine.

The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

#### Synthetic Polycationic Agents

Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin<sup>®</sup>, and lipofectAMINE<sup>®</sup> are monomers that form polycationic complexes when combined with polynucleotides or polypeptides.

#### Immunodiagnostic Assays

Neisserial antigens of the invention can be used in immunoassays to detect antibody levels (or, conversely, anti-Neisserial antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to Neisserial proteins within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are

assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, *etc.*) required for the conduct of the assay, as well as suitable set of assay instructions.

### **Nucleic Acid Hybridisation**

“Hybridization” refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook *et al.* [*supra*] Volume 2, chapter 9, pages 9.47 to 9.57.

“Stringency” refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200°C below the calculated  $T_m$  of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1 µg for a plasmid or phage digest to  $10^{-9}$  to  $10^{-8}$  g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy



yeast gene can be detected with an exposure time of only 1 hour starting with 1  $\mu$ g of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of  $10^8$  cpm/ $\mu$ g. For a single-copy mammalian gene a conservative approach would start with 10  $\mu$ g of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than  $10^8$  cpm/ $\mu$ g, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature ( $T_m$ ) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

$$T_m = 81 + 16.6(\log_{10} C_i) + 0.4[\%(G + C)] - 0.6(\%\text{formamide}) - 600/n - 1.5(\%\text{mismatch}).$$

where  $C_i$  is the salt concentration (monovalent ions) and  $n$  is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) *Anal. Biochem.* 138: 267-284).

In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (*ie.* stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

In general, convenient hybridization temperatures in the presence of 50% formamide are 42°C for a probe with is 95% to 100% homologous to the target fragment, 37°C for 90% to 95% homology, and 32°C for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be

washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

### **Nucleic Acid Probe Assays**

Methods such as PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes according to the invention can determine the presence of cDNA or mRNA. A probe is said to "hybridize" with a sequence of the invention if it can form a duplex or double stranded complex, which is stable enough to be detected.

The nucleic acid probes will hybridize to the Neisserial nucleotide sequences of the invention (including both sense and antisense strands). Though many different nucleotide sequences will encode the amino acid sequence, the native Neisserial sequence is preferred because it is the actual sequence present in cells. mRNA represents a coding sequence and so a probe should be complementary to the coding sequence; single-stranded cDNA is complementary to mRNA, and so a cDNA probe should be complementary to the non-coding sequence.

The probe sequence need not be identical to the Neisserial sequence (or its complement) — some variation in the sequence and length can lead to increased assay sensitivity if the nucleic acid probe can form a duplex with target nucleotides, which can be detected. Also, the nucleic acid probe can include additional nucleotides to stabilize the formed duplex. Additional Neisserial sequence may also be helpful as a label to detect the formed duplex. For example, a non-complementary nucleotide sequence may be attached to the 5' end of the probe, with the remainder of the probe sequence being complementary to a Neisserial sequence. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the a Neisserial sequence in order to hybridize therewith and thereby form a duplex which can be detected.

The exact length and sequence of the probe will depend on the hybridization conditions, such as temperature, salt condition and the like. For example, for diagnostic applications, depending on the complexity of the analyte sequence, the nucleic acid probe typically contains at least 10-20 nucleotides, preferably 15-25, and more preferably at least 30 nucleotides, although it may be shorter than this. Short primers generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.

Probes may be produced by synthetic procedures, such as the triester method of Matteucci *et al.* [*J. Am. Chem. Soc.* (1981) 103:3185], or according to Urdea *et al.* [*Proc. Natl. Acad. Sci. USA* (1983) 80: 7461], or using commercially available automated oligonucleotide synthesizers.

The chemical nature of the probe can be selected according to preference. For certain applications, DNA or RNA are appropriate. For other applications, modifications may be incorporated *eg.* backbone modifications, such as phosphorothioates or methylphosphonates, can be used to increase *in vivo* half-life, alter RNA affinity, increase nuclease resistance *etc.* [*eg.* see Agrawal & Iyer (1995) *Curr Opin Biotechnol* 6:12-19; Agrawal (1996) *TIBTECH* 14:376-387]; analogues such as peptide nucleic acids may also be used [*eg.* see Corey (1997) *TIBTECH* 15:224-229; Buchardt *et al.* (1993) *TIBTECH* 11:384-386].

One example of a nucleotide hybridization assay is described by Urdea *et al.* in international patent application WO92/02526 [see also US patent 5,124,246].

Alternatively, the polymerase chain reaction (PCR) is another well-known means for detecting small amounts of target nucleic acids. The assay is described in: Mullis *et al.* [*Meth. Enzymol.* (1987) 155: 335-350]; US patent 4,683,195; and US patent 4,683,202. Two "primer" nucleotides hybridize with the target nucleic acids and are used to prime the reaction. The primers can comprise sequence that does not hybridize to the sequence of the amplification target (or its complement) to aid with duplex stability or, for example, to incorporate a convenient restriction site. Typically, such sequence will flank the desired Neisserial sequence.

A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a threshold amount of target nucleic acids are generated by the polymerase, they can be detected by more traditional methods, such as Southern blots. When using the Southern blot method, the labelled probe will hybridize to the Neisserial sequence (or its complement).

Also, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook *et al* [*supra*]. mRNA, or cDNA generated from mRNA using a polymerase enzyme, can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labelled probe and then washed to remove any unhybridized probe. Next, the duplexes

containing the labeled probe are detected. Typically, the probe is labelled with a radioactive moiety.

## EXAMPLES

The examples describe nucleic acid sequences which have been identified in *N. meningitidis*, and *N. gonorrhoeae* along with their respective and putative translation products. Not all of the nucleic acid sequences are complete *ie.* they encode less than the full-length wild-type protein.

The examples are generally in the following format:

- a nucleotide sequence which has been identified in *N. meningitidis*
- the putative translation product of said *N. meningitidis* sequence
- a computer analysis of said translation product based on database comparisons
- a corresponding nucleotide sequence identified from *N. gonorrhoeae*
- the putative translation product of said *N. gonorrhoeae* sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a corresponding nucleotide sequence identified from strain A of *N. meningitidis*
- the putative translation product of said *N. meningitidis* strain A sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a description of the characteristics of the protein which indicates that it might be suitably antigenic or immunogenic.

Sequence comparisons were performed at NCBI (<http://www.ncbi.nlm.nih.gov>) using the algorithms BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx, & tBLASTx [eg. see also Altschul *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:2289-3402]. Searches were performed against the following databases: non-redundant GenBank+EMBL+DDBJ+PDB sequences and non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR sequences.

Dots within nucleotide sequences represent nucleotides which have been arbitrarily introduced in order to maintain a reading frame. In the same way, double-underlined nucleotides were removed. Lower case letters represent ambiguities which arose during

alignment of independent sequencing reactions (some of the nucleotide sequences in the examples are derived from combining the results of two or more experiments).

Nucleotide sequences were scanned in all six reading frames to predict the presence of hydrophobic domains using an algorithm based on the statistical studies of Esposti *et al.* [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. These domains represent potential transmembrane regions or hydrophobic leader sequences.

Open reading frames were predicted from fragmented nucleotide sequences using the program ORFFINDER (NCBI).

Underlined amino acid sequences indicate possible transmembrane domains or leader sequences in the ORFs, as predicted by the PSORT algorithm (<http://www.psort.nibb.ac.jp>). Functional domains were also predicted using the MOTIFS program (GCG Wisconsin & PROSITE).

For each of the following examples: based on the presence of a putative leader sequence and/or several putative transmembrane domains (single-underlined) in the gonococcal protein, it is predicted that the proteins from *N. meningitidis* and *N. gonorrhoeae*, and their respective epitopes, could be useful antigens or immunogenic compositions for vaccines or diagnostics.

The standard techniques and procedures which may be employed in order to perform the invention (*e.g.* to utilize the disclosed sequences for vaccination or diagnostic purposes) were summarized above. This summary is not a limitation on the invention but, rather, gives examples that may be used, but are not required.

In particular, the following methods were used to express, purify and biochemically characterize the proteins of the invention.

#### **Chromosomal DNA Preparation**

*N.meningitidis* strain 2996 was grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20%(w/v) Sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl<sub>3</sub>/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation.

The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading the OD at 260 nm.

### Oligonucleotide design

Synthetic oligonucleotide primers were designed on the basis of the coding sequence of each ORF, using (a) the meningococcus B sequence when available, or (b) the gonococcus/meningococcus A sequence, adapted to the codon preference usage of meningococcus as necessary. Any predicted signal peptides were omitted, by designing the 5' primers to sequence immediately downstream from the predicted leader sequence.

For most ORFs, the 5' primers included two restriction enzyme recognition sites (*Bam*HI-*Nde*I, *Bam*HI-*Nhe*I, *Eco*RI-*Nde*I or *Eco*RI-*Nhe*I), depending on the restriction pattern of the gene of interest. The 3' primers included a *Xho*I or a *Hind*III restriction site (table 1). This procedure was established in order to direct the cloning of each amplification product (corresponding to each ORF) into two different expression systems: pGEX-KG (using *Bam*HI-*Xho*I, *Bam*HI-*Hind*III, *Eco*RI-*Xho*I or *Eco*RI-*Hind*III), and pET21b+ (using *Nde*I-*Xho*I, *Nhe*I-*Xho*I, *Nde*I-*Hind*III or *Nhe*I-*Hind*III).

5'-end primer tail:	<u>CGCGGATCCCATATG</u>	( <i>Bam</i> HI- <i>Nde</i> I )
	<u>CGCGGATCCGCTAGC</u>	( <i>Bam</i> HI- <i>Nhe</i> I)
	<u>CCGGAATTCTACATATG</u>	( <i>Eco</i> RI- <i>Nde</i> I)
	<u>CCGGAATTCTAGCTAGC</u>	( <i>Eco</i> RI- <i>Nhe</i> I)
3'-end primer tail:	<u>CCCGCTCGAG</u>	( <i>Xho</i> I)
	<u>CCCGCTCGAG</u>	( <i>Hind</i> III)

For cloning ORFs into the pGEX-His vector, the 5' and 3' primers contained only one restriction enzyme site (*Eco*RI, *Kpn*I or *Sal*I for the 5' primers and *Pst*I, *Xba*I, *Sph*I or *Sal*I for the 3' primers). Again restriction sites were chosen according to the particular restriction pattern of the gene (table 1).

5'-end primer tail:	(AAA) <u>AAAGAATTC</u>	( <i>Eco</i> RI )
	(AAA) <u>AAAGGTACC</u>	( <i>Kpn</i> I)
3'-end primer tail:	(AAA) <u>AAACTGCAG</u>	( <i>Pst</i> I)
	(AAA) <u>AAATCTAGA</u>	( <i>Xba</i> I)

5' or 3'-end primer tail:                      AAAGCATGC                      (*Sph*I)  
    AAAAAAGTCGAC                      (*Sal*I)

As well as containing the restriction enzyme recognition sequences, the primers included nucleotides which hybridized to the sequence to be amplified. The melting temperature depended on the number and type of hybridising nucleotides in the whole primer, and was determined for each primer using the formulae:

$$T_m = 4 (G+C) + 2 (A+T) \quad (\text{tail excluded})$$

$$T_m = 64.9 + 0.41 (\% \text{ GC}) - 600/N \quad (\text{whole primer})$$

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-55°C for the hybridising region alone.

Table 1 shows the forward and reverse primers used for each amplification. In certain cases, the sequence of the primer does not exactly match the sequence of the predicted ORF. This is because when initial amplifications were performed, the complete 5' and/or 3' sequences for some meningococcal B ORFs were not known. However the corresponding sequences had been identified in *Gonococcus* or in *Meningococcus A*. Hence, when the *Meningococcus B* sequence was incomplete or uncertain, *Gonococcal* or *Meningococcal A* sequences were used as the basis for primer design. These sequences were altered to take account of codon preference. It can be appreciated that, once the complete sequence is identified, this approach will no longer be necessary.

Oligonucleotides were synthesized using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH<sub>4</sub>OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in either 100µl or 1.0ml of water. The OD<sub>260</sub> was determined using a Perkin Elmer Lambda Bio spectrophotometer and the concentration adjusted to 2-10pmol/µl.

### Amplification

The standard PCR protocol was as follows: 50-200ng of genomic DNA was used as a template in the presence of 20-40µM of each oligonucleotide primer, 400-800µM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl<sub>2</sub>), 2.5 units *TaqI* DNA polymerase (using

Perkin-Elmer AmpliTaq, GIBCO Platinum, Pwo DNA polymerase, or Tahara Shuzo Taq polymerase). In some cases, PCR was optimised by the addition of 10µl DMSO or 50µl 2M Betaine.

After a hot start (adding the polymerase during a preliminary 3 minute incubation of the whole mix at 95°C), each sample underwent a two-step amplification. The first 5 cycles were performed using the hybridization temperature that excluded the restriction enzyme tail of the primer (see above). This was followed by 30 cycles using the hybridization temperature calculated for the whole length oligos. The cycles were completed with a 10 minute extension step at 72°C. The standard cycles were as follows:

	<b>Denaturation</b>	<b>Hybridisation</b>	<b>Elongation</b>
First 5 cycles	30 seconds 95°C	30 seconds 50-55°C	30-60 seconds 72°C
Last 30 cycles	30 seconds 95°C	30 seconds 65-70°C	30-60 seconds 72°C

Elongation times varied according to the length of the ORF to be amplified. Amplifications were performed using either a 9600 or a 2400 Perkin Elmer GeneAmp PCR System. To check the results, 1/10 of the amplification volume was loaded onto a 1-1.5% (w/v) agarose gel and the size of each amplified fragment compared with a DNA molecular weight marker.

The amplified DNA was either loaded directly on a 1% agarose gel or first precipitated with ethanol and resuspended in a volume suitable to be loaded on a 1.0% agarose gel. The DNA fragment corresponding to the band of correct size was purified using the Qiagen Gel Extraction Kit, following the manufacturer's protocol. DNA fragments were eluted in a volume of 30µl or 50µl with either H<sub>2</sub>O or 10mM Tris, pH 8.5.

### **Digestion of PCR fragments**

The purified DNA corresponding to the amplified fragment was doubly-digested with the appropriate restriction enzymes for; cloning into pET-21b+ and expressing the protein as a C-terminus His-tagged fusion, for cloning into pGEX-KG and expressing the protein as a N-



terminus GST-fusion, and for cloning into pGEX-His and expressing the protein as a N-terminus GST-His tagged fusion.

Each purified DNA fragment was incubated at 37°C for 3 hours to overnight with 20 units of appropriate restriction enzyme (New England Biolabs) in a volume of either 30 or 40 µl in the presence of suitable digestion buffer. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted in a volume of 30 µl or 50 µl with either H<sub>2</sub>O or 10mM Tris, pH 8.5. The DNA concentration was determined by quantitative agarose gel electrophoresis (1.0% gel) in the presence of a titrated molecular weight marker.

#### **Digestion of the cloning vectors (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG, and pGEX-His)**

The vector pGEX-His is a modified pGEX-2T vector carrying a region encoding six histidine residues upstream of the thrombin cleavage site and containing the multiple cloning site of the vector pTRC99 (Pharmacia). 10 µg plasmid was double-digested with 50 units of each restriction enzyme in 200 µl reaction volume in the presence of appropriate buffer by overnight incubation at 37°C. After loading the whole digestion on a 1% agarose gel, the band corresponding to the digested vector was purified from the gel using the Qiagen QIAquick Gel Extraction Kit and the DNA was eluted in 50 µl of 10 mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD<sub>260</sub> of the sample, and adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

10 µg of plasmid vector was doubly-digested with 50 units of each restriction enzyme in a volume of 200 µl with the appropriate buffer overnight at 37°C. The digest was loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit. DNA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD<sub>260nm</sub> and the concentration adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

#### **Cloning**

For some ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated in both pET22b and pGEX-KG. In a final volume of 20 µl, a molar

ratio of 3:1 fragment/vector was ligated using 0.5  $\mu$ l of NEB T4 DNA ligase (400 units/ $\mu$ l), in the presence of the buffer supplied by the manufacturer. The reaction was incubated at room temperature for 3 hours. In some experiments, ligation was performed using the Boehringer "Rapid Ligation Kit", following the manufacturer's instructions.

In order to introduce the recombinant plasmid in a suitable strain, 100  $\mu$ l *E. coli* DH5 competent cells were incubated with the ligase reaction solution for 40 minutes on ice, then at 37°C for 3 minutes, then, after adding 800  $\mu$ l LB broth, again at 37°C for 20 minutes. The cells were then centrifuged at maximum speed in an Eppendorf microfuge and resuspended in approximately 200  $\mu$ l of the supernatant. The suspension was then plated on LB ampicillin (100 mg/ml ).

The screening of the recombinant clones was performed by growing 5 randomly-chosen colonies overnight at 37 °C in either 2 ml (pGEX or pTC clones) or 5ml (pET clones) LB broth + 100  $\mu$ g/ml ampicillin. The cells were then pelleted and the DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions, to a final volume of 30  $\mu$ l. 5  $\mu$ l of each individual miniprep (approximately 1g ) were digested with either *NdeI/XhoI* or *BamHI/XhoI* and the whole digestion loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1Kb DNA Ladder, GIBCO). The screening of the positive clones was made on the base of the correct insert size.

For other ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated into both pET21b+ and pGEX-KG. A molar ratio of of 3:1 fragment/vector was used in a final volume of 20 $\mu$ l, that included 0.5 $\mu$ l T4 DNA ligase (400 units/ $\mu$ l, NEB) and ligation buffer supplied by the manufacturer. The reaction was performed at room temperature for 3 hours. In some experiments, ligation was performed using the Boehringer "Rapid Ligation Kit" and the manufacturer's protocol.

Recombinant plasmid was transformed into 100 $\mu$ l of competent *E. coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice then at 37°C for 3 minutes. This was followed by the addition of 800 $\mu$ l LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200 $\mu$ l of the supernatant and plated onto LB ampicillin (100mg/ml ) agar.

Screening for recombinant clones was performed by growing 5 randomly selected colonies overnight at 37°C in either 2.0ml (pGEX-KG clones) or 5.0ml (pET clones) LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

ORFs were cloned into PGEX-His, by doubly-digesting the PCR product and ligating into similarly digested vector. After cloning, recombinant plasmids were transformed into the *E.coli* host W3110. Individual clones were grown overnight at 37°C in LB broth with 50µg/ml ampicillin.

Certain ORFs may be cloned into the pGEX-HIS vector using *EcoRI-PstI* cloning sites, or *EcoRI-SalI*, or *SalI-PstI*. After cloning, the recombinant plasmids may be introduced in the *E.coli* host W3110.

## Expression

Each ORF cloned into the expression vector may then be transformed into the strain suitable for expression of the recombinant protein product. 1 µl of each construct was used to transform 30 µl of *E.coli* BL21 (pGEX vector), *E.coli* TOP 10 (pTRC vector) or *E.coli* BL21-DE3 (pET vector), as described above. In the case of the pGEX-His vector, the same *E.coli* strain (W3110) was used for initial cloning and expression. Single recombinant colonies were inoculated into 2ml LB+Amp (100 µg/ml), incubated at 37°C overnight, then diluted 1:30 in 20 ml of LB+Amp (100 µg/ml) in 100 ml flasks, making sure that the OD<sub>600</sub> ranged between 0.1 and 0.15. The flasks were incubated at 30°C into gyratory water bath shakers until OD indicated exponential growth suitable for induction of expression (0.4-0.8 OD for pET and pTRC vectors; 0.8-1 OD for pGEX and pGEX-His vectors). For the pET, pTRC and pGEX-His vectors, the protein expression was induced by addition of 1mM IPTG, whereas in the case of pGEX system the final concentration of IPTG was 0.2 mM. After 3 hours incubation at 30°C, the final concentration of the sample was checked by OD. In order to check expression, 1ml of each sample was removed, centrifuged in a microfuge, the pellet

resuspended in PBS, and analysed by 12% SDS-PAGE with Coomassie Blue staining. The whole sample was centrifuged at 6000g and the pellet resuspended in PBS for further use.

#### **GST-fusion proteins large-scale purification.**

For some ORFs, a single colony was grown overnight at 37°C on LB+Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600 ml of fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD<sub>550</sub> 0.8-1. Protein expression was induced with 0.2mM IPTG followed by three hours incubation. The culture was centrifuged at 8000 rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml cold PBS. The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia) (previously washed with PBS) and incubated at room temperature for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold PBS for 10 minutes, resuspended in 1ml cold PBS, and loaded on a disposable column. The resin was washed twice with 2ml cold PBS until the flow-through reached OD<sub>280</sub> of 0.02-0.06. The GST-fusion protein was eluted by addition of 700µl cold Glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl) and fractions collected until the OD<sub>280</sub> was 0.1. 21µl of each fraction were loaded on a 12% SDS gel using either Biorad SDS-PAGE Molecular weight standard broad range (M1) (200, 116.25, 97.4, 66.2, 45, 31, 21.5, 14.4, 6.5 kDa) or Amersham Rainbow Marker (M'') (220, 66, 46, 30, 21.5, 14.3 kDa) as standards. As the MW of GST is 26kDa, this value must be added to the MW of each GST-fusion protein.

For other ORFs, for each clone to be purified as a GST-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp. (100µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD<sub>550nm</sub> reached 0.6-0.8. Recombinant protein expression was induced by addition of IPTG (final concentration 0.2mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml cold PBS. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia), previously equilibrated with PBS, and incubated at room temperature with gentle agitation for 30 min. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batchwise) with 10ml cold PBS for 10 min, resuspended in 1ml cold PBS, and loaded onto a disposable column. The resin continued to be washed with cold PBS, until the OD<sub>280nm</sub> of the flow-through reached 0.02-0.01. The GST-fusion protein was eluted by addition of 700µl cold glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl pH 8.0) and fractions collected, until the OD<sub>280nm</sub> of the eluate indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. The molecular mass of the purified proteins was determined using either the Bio-Rad broad range molecular weight standard (M1) (200, 116, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa) or the Amersham Rainbow Marker (M2) (220, 66.2, 46.0, 30.0, 21.5, 14.3 kDa). The molecular weights of GST-fusion proteins are a combination of the 26 kDa GST protein and its fusion partner. Protein concentrations were estimated using the Bradford assay.

#### **His-fusion soluble proteins large-scale purification.**

For some ORFs, a single colony was grown overnight at 37°C on a LB + Amp agar plate. The bacteria were inoculated into 20ml of LB+Amp liquid culture and incubated overnight in a water bath shaker. Bacteria were diluted 1:30 into 600ml fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD<sub>550</sub> 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000 rpm at 4°C, the supernatant was discarded and the bacterial pellet was resuspended in 7.5ml cold 10mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Ni<sup>2+</sup>-resin (Pharmacia) (previously washed with 10mM imidazole buffer) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold 10mM imidazole buffer for 10 minutes,

resuspended in 1ml cold 10mM imidazole buffer and loaded on a disposable column. The resin was washed at 4°C with 2ml cold 10mM imidazole buffer until the flow-through reached the O.D<sub>280</sub> of 0.02-0.06. The resin was washed with 2ml cold 20mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 20 mM imidazole, pH 8) until the flow-through reached the O.D<sub>280</sub> of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl cold 250mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 250 mM imidazole, pH 8) and fractions collected until the O.D<sub>280</sub> was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

#### **His-fusion insoluble proteins large-scale purification.**

A single colony was grown overnight at 37 °C on a LB + Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600ml fresh medium and let to grow at the optimal temperature (37°C) to O.D<sub>550</sub> 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml buffer B (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed twice and centrifuged again. The supernatant was stored at -20°C, while the pellets were resuspended in 2 ml guanidine buffer (6M guanidine hydrochloride, 100mM phosphate buffer, 10 mM Tris-HCl, pH 7.5) and treated in a homogenizer for 10 cycles. The product was centrifuged at 13000 rpm for 40 minutes. The supernatant was mixed with 150µl Ni<sup>2+</sup>-resin (Pharmacia) (previously washed with buffer B) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700 g for 5 minutes at 4°C. The resin was washed twice with 10 ml buffer B for 10 minutes, resuspended in 1ml buffer B, and loaded on a disposable column. The resin was washed at room temperature with 2ml buffer B until the flow-through reached the OD<sub>280</sub> of 0.02-0.06. The resin was washed with 2ml buffer C (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the flow-through reached the O.D<sub>280</sub> of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl elution buffer (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions collected until the OD<sub>280</sub> was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

**Purification of His-fusion proteins.**

For each clone to be purified as a His-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp (100 µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD<sub>550nm</sub> reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml of either (i) cold buffer A (300mM NaCl, 50mM phosphate buffer, 10mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8) for insoluble proteins. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6M guanidine hydrochloride, 100mM phosphate buffer, 10mM Tris-HCl, pH 7.5) and treated with a Dounce homogenizer for 10 cycles. The homogenate was centrifuged at 13 000xg for 40 min and the supernatant retained.

Supernatants for both soluble and insoluble preparations were mixed with 150µl Ni<sup>2+</sup>-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to manufacturers protocol. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD<sub>280nm</sub> of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the the OD<sub>280nm</sub> of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700µl of either (i) cold elution buffer A (300mM NaCl, 50mM phosphate buffer, 250mM imidazole, pH 8.0) or (ii) elution buffer B (8 M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions

collected until the  $O.D_{280nm}$  indicated all the recombinant protein was obtained. 20 $\mu$ l aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. Protein concentrations were estimated using the Bradford assay.

### **His-fusion proteins renaturation**

In the cases where denaturation was required to solubilize proteins, a renaturation step was employed prior to immunization. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10%(v/v). The proteins were diluted to 200 $\mu$ g/ml using dialysis buffer I (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Alternatively, 10% glycerol was added to the denatured proteins. The proteins were then diluted to 20 $\mu$ g/ml using dialysis buffer I (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, 2M urea, pH 8.8) and dialysed against the same buffer at 4°C for 12-14 hours. The protein was further dialysed against dialysis buffer II (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Protein concentration was evaluated using the formula:

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

### **Purification of proteins**

To analyse the solubility, pellets obtained from 3.0ml cultures were resuspended in 500 $\mu$ l buffer M1 (PBS pH 7.2). 25 $\mu$ l of lysozyme (10mg/ml) was added and the bacteria incubated for 15 min at 4°C. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and the pellet resuspended in buffer M2 [8M urea, 0.5M NaCl, 20mM imidazole and 0.1M  $\text{NaH}_2\text{PO}_4$ ] and incubated for 3 to 4 hours at 4°C. After centrifugation, the supernatant was collected and the pellet resuspended in buffer M3 [6M guanidinium-HCl, 0.5M NaCl, 20mM imidazole and 0.1M  $\text{NaH}_2\text{PO}_4$ ] overnight at 4°C. The



supernatants from all steps were analysed by SDS-PAGE. Some proteins were found to be soluble in PBS, others need urea or guanidium-HCl for solubilization.

For preparative scale purifications, 500ml cultures were induced and fusion proteins solubilized in either buffer M1, M2 or M3 using the procedure described above. Crude extracts were loaded onto a Ni-NTA superflow column (Quiagen) equilibrated with buffer M1, M2 or M3 depending on the solubilization buffer employed. Unbound material was eluted by washing the column with the same buffer. The recombinant fusion protein was eluted with the corresponding buffer containing 500mM imidazole then dialysed against the same buffer in the absence of imidazole.

### **Mice immunisations**

20µg of each purified protein are used to immunise mice intraperitoneally. In the case of some ORFs, Balb-C mice were immunised with Al(OH)<sub>3</sub> as adjuvant on days 1, 21 and 42, and immune response was monitored in samples taken on day 56. For other ORFs, CD1 mice could be immunised using the same protocol. For ORFs 25 and 40, CD1 mice were immunised using Freund's adjuvant, and the same immunisation protocol was used, except that the immune response was measured on day 42, rather than 56. Similarly, for still other ORFs, CD1 mice were immunised with Freund's adjuvant, but the immune response was measured on day 49. Alternatively, 20µg of each purified protein was mixed with Freund's adjuvant and used to immunise CD1 mice intraperitoneally. For many of the proteins, the immunization was performed on days 1, 21 and 35, and immune response was monitored in samples taken on days 34 and 49. For some proteins, the third immunization was performed on day 28, rather than 35, and the immune response was measured on days 20 and 42, rather than 34 and 49.

### **ELISA assay (sera analysis)**

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 7ml of Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10000 rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated

for 2 hours at room temperature and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200 µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200 µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN<sub>3</sub> in PBS) were added to each well and the plates incubated for 90 minutes at 37°C. Wells were washed three times with PBT. 100 µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 µl of substrate buffer for HRP (25 ml of citrate buffer pH5, 10 mg of O-phenildiamine and 10 µl of H<sub>2</sub>O) were added to each well and the plates were left at room temperature for 20 minutes. 100 µl H<sub>2</sub>SO<sub>4</sub> was added to each well and OD<sub>490</sub> was followed. The ELISA was considered positive when OD490 was 2.5 times the respective pre-immune sera.

Alternatively, The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10 000rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN<sub>3</sub> in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-

phenildiamine and 10µl of H<sub>2</sub>O<sub>2</sub>) were added to each well and the plates were left at room temperature for 20 minutes. 100µl of 12.5% H<sub>2</sub>SO<sub>4</sub> was added to each well and OD<sub>490</sub> was followed. The ELISA titers were calculated arbitrarily as the dilution of sera which gave an OD<sub>490</sub> value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD<sub>490</sub> of 0.4 was higher than 1:400.

#### **FACScan bacteria Binding Assay procedure.**

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN<sub>3</sub>) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD<sub>620</sub> of 0.07. 100µl bacterial cells were added to each well of a Costar 96 well plate. 100µl of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200µl/well of blocking buffer in each well. 100µl of R-Phicoerytrin conjugated F(ab)<sub>2</sub> goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200µl/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200µl/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

#### **OMV preparations**

Bacteria were grown overnight on 5 GC plates, harvested with a loop and resuspended in 10 ml 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes and the bacteria disrupted by sonication for 10' on ice ( 50% duty cycle, 50% output ). Unbroken cells were removed by centrifugation at 5000g for 10 minutes and the total cell envelope

fraction recovered by centrifugation at 50000g at 4°C for 75 minutes. To extract cytoplasmic membrane proteins from the crude outer membranes, the whole fraction was resuspended in 2% sarkosyl (Sigma) and incubated at room temperature for 20 minutes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, and the supernatant further ultracentrifuged at 50000g for 75 minutes to pellet the outer membranes. The outer membranes were resuspended in 10mM Tris-HCl, pH8 and the protein concentration measured by the Bio-Rad Protein assay, using BSA as a standard.

### **Whole Extracts preparation**

Bacteria were grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30' minutes.

### **Western blotting**

Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3 % Tris base, 1.44 % glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

### **Bactericidal assay**

MC58 and 2996 strains were grown overnight at 37°C on chocolate agar plates. 5-7 colonies were collected and used to inoculate 7ml Mueller-Hinton broth. The suspension was incubated at 37°C on a nutator and let to grow until OD<sub>620</sub> was in between 0.5-0.8. The culture was aliquoted into sterile 1.5ml Eppendorf tubes and centrifuged for 20 minutes at maximum speed in a microfuge. The pellet was washed once in Gey's buffer (Gibco) and resuspended in the same buffer to an OD<sub>620</sub> of 0.5, diluted 1:20000 in Gey's buffer and stored at 25°C.

50µl of Gey's buffer/1% BSA was added to each well of a 96-well tissue culture plate. 25µl of diluted (1:100) mice sera (dilution buffer: Gey's buffer/0.2% BSA) were added to each well and the plate incubated at 4°C. 25µl of the previously described bacterial suspension were added to each well. 25µl of either heat-inactivated (56°C waterbath for 30 minutes) or normal baby rabbit complement were added to each well. Immediately after the addition of the baby rabbit complement, 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 0). The 96-well plate was incubated for 1 hour at 37°C with rotation and then 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 1). After overnight incubation the colonies corresponding to time 0 and time 1h were counted.

### Gene Variability

The *ORF4* and *919* genes were amplified by PCR on chromosomal DNA extracted from various *Neisseria* strains (see list of strains). The following oligonucleotides used as PCR primers were designed in the upstream and downstream regions of the genes:

orf 4.1	(forward)	CGAATCCGGACGGCAGGACTC
orf 4.3	(reverse)	GGCAGGGAATGGCGGATTAAAG
919.1	(forward)	AAAATGCCTCTCCACGGCTG or CTGCGCCCTGTGTTAAAATCCCCCT
919.6	(reverse)	CAAATAAGAAAGGAATTTTG or GGTATCGCAAAACTTCGCCTTAATGCG

The PCR cycling conditions were:

1 cycle	2 min. at 94°
30 cycles	30 sec. at 94°
	30 sec. at ~ 54° or ~ 60° ( in according to Tm of the primers)
	40 sec. at 72°
1 cycle	7 min. at 72°

The PCR products were purified from 1 % agarose gel and sequenced using the following primers:

orf 4.1 (forward)	CGAATCCGGACGGCAGGACTC
orf 4.2 (forward)	CGACCGCGCCTTTGGGACTG
orf 4.3 (reverse)	GGCAGGGAATGGCGGATTAAAG
orf 4.4 (reverse)	TCTTTGAGTTTGATCCAACC

- 919.1 (forward) AAAATGCCTCTCCACGGCTG or  
CTGCGCCCTGTGTTAAAATCCCCCT
- 919.2 (forward) ATCCTTCCGCCTCGGCTGCG
- 919.3 (forward) AAAACAGCGGCACAATCGAC
- 919.4 (forward) ATAAGGGCTACCTCAAATC
- 919.5 (forward) GCGCGTGGATTATTTTGGG
- 919.6 (reverse) CAAATAAGAAAGGAATTTTG or  
GGTATCGCAAACTTCGCCTTAATGCG
- 919.7 (reverse) CCCAAGGTAATGTAGTGCCG
- 919.8 (reverse) TAAAAAAAAGTTCGACAGGG
- 919.9 (reverse) CCGTCCGCCTGTCGTCGCCC
- 919.10 (reverse) TCGTTCCGGCGGGGTCGGGG

All documents cited herein are incorporated by reference in their entireties.

The following Examples are presented to illustrate, not limit, the invention

#### EXAMPLE 1

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

**Table 1: Oligonucleotides used for PCR for Examples 2-10**

ORF	Primer	Sequence	Restriction sites
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT <SEQ ID 3021>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA <SEQ ID 3022>	XhoI
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA <SEQ ID 3023>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC <SEQ ID 3024>	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT <SEQ ID 3025>	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC <SEQ ID 3026>	XhoI
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC <SEQ ID 3027>	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG <SEQ ID 3028>	XhoI
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT <SEQ ID 3029>	BamHI-NdeI

128	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC <SEQ ID 3030>	XhoI
	Forward	CGCGGATCCCATATG-ACTGACAACGCACT <SEQ ID 3031>	BamHI-NdeI
206	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA <SEQ ID 3032>	XhoI
	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA <SEQ ID 3033>	BamHI-NdeI
287	Reverse	CCCGCTCGAG-TTCTGTAAAAAAGTATGTGC <SEQ ID 3034>	XhoI
	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG <SEQ ID 3035>	EcoRI-NheI
406	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTGGCC <SEQ ID 3036>	XhoI
	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG <SEQ ID 3037>	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG <SEQ ID 3038>	XhoI

### Localization of the ORFs

The following DNA and amino acid sequences are identified by titles of the following form: [g, m, or a] [#].[seq or pep], where "g" means a sequence from *N. gonorrhoeae*, "m" means a sequence from *N. meningitidis B*, and "a" means a sequence from *N. meningitidis A*; "#" means the number of the sequence; "seq" means a DNA sequence, and "pep" means an amino acid sequence. For example, "g001.seq" refers to an *N. gonorrhoeae* DNA sequence, number 1. The presence of the suffix "-1" to these sequences indicates an additional sequence found for the same ORF, thus, data for an ORF having both an unsuffixed and a suffixed sequence designation applies to both such designated sequences. Further, open reading frames are identified as ORF #, where "#" means the number of the ORF, corresponding to the number of the sequence which encodes the ORF, and the ORF designations may be suffixed with ".ng" or ".a", indicating that the ORF corresponds to a *N. gonorrhoeae* sequence or a *N. meningitidis A* sequence, respectively. The word "partial" before a sequence indicates that the sequence may be a partial or a complete ORF. Computer analysis was performed for the comparisons that follow between "g", "m", and "a" peptide sequences; and therein the "pep" suffix is implied where not expressly stated. Further, in the event of a conflict between the text immediately preceding and describing which sequences are being compared, and the

279 gnm4.seq

1	ATAACGCGGA	TTTGCGGCTG	CTTGATTTC	ACGGTTTTCA	GGGCTTCGGC
51	AAGTTTGTCTG	GCGGCGGGTT	TCATCAGGCT	GCAATGGGAA	GGTACGGACA
101	CGGGCAGCGG	GAGGGCGCGT	TTGGCGCGCA	CTTCTTTTGC	GGCAGCCATG
151	GCGCGTCCGA	CAGGCGGCGC	CTTGCCCTCG	ATCAGGATTT	GTCCGGGTGA
201	GTTGAAGTTG	ACGGCTTCGA	CCACTTCGCT	TTGGCGGGT	CTGGACAAAA
251	TGGCTTTAAC	CTGCTCATCT	TCCAAGCCGA	GATCGCCGC	CATTGCGCCC
301	ACGCCTTGCG	GTACGGCGGA	CTGCATCAGT	TCGGCGCGCA	GGCGCACGAG
351	TTTGACCGCG	TCGGCAAAAT	TCAATGCGCC	GGCGCAAT	AATGCGGTGT
401	ATTCGCCGAG	GCTGTGTCCG	GCAACGGCGG	CAGGCGTTAT	GCCGCCCGCT
451	TCTAAATAG				

```

1  ITRICGCLIS TVFRASASLS AAGFIRLQWE GTDTGSGRAR LAPASLAAAM
51  ARPTAAALPA ITICPGELKL TASTSLWAA SAQMALTCSS SKPRIAAIAP
101 TPCGTADCIS SARRRTSLTA SAKFNAPAAAT SAVYSPRLCP ATAAGVLPPA
151 SK*

```

1	atgacgcgga	tttgcggtctg	cttgatttca	acggttttga	gtgttttcggc
51	aagtttgtcg	gcggcggggt	tcatacggct	gcaatgggaa	ggaacggata
101	ccggcagcga	caggcgcgct	ttgttcctgga	cttcttttggc	ggcagccatg
151	gtgcgctcga	cggcgggcgc	ttgctctcga	atcacgactt	gtccggggcga
201	gttgaagtgtg	acggcttcga	ccacttcgcc	ctgtgcggaat	tcggcacaaa
251	tctgcctgac	ctgttcacat	tccaaaccca	aaatggccgc	cattgcccct
301	acgccttgcg	gtacggcgga	ctgcatcagt	tcggcgcgca	ggcggacgag
351	tttgacggga	tcggcaaaat	ccaatgcttc	ggcggcgaca	agcgcggtgt
401	attcccgag	gctgtgtccg	gcaacggcgg	caggcgtttt	gccgcccaact
451	tccaaaataq				

```

1  MTRICGLIS TVLSVSASLS AAGFIRLQWE GTDTGSGRAR LAPASLAAAM
51  VRPTAAALPA ITTCPGELKL TASTTSPCAD SAQICLTSS SKPKMAAIAP
101 TPCGTADCIS SARRRTSLTA SAKSNASAAT SAVISPRLCP ATAAGVLPP
151 SK*

```

```

              10      20      30      40      50      60
m279.pep    ITRICGCLISTVFRASASLSAAGFIRLOWEGTDTGSGRARLAPASLAAMARPTAAALPA
             :|||:|||::: : |||:|||::: : |||:|||::: : |||:|||::: : |||:|||:::
g279        MTRICGCLISTVLVSASLSAAGFIRLOWEGTDTGSGRARLAPASLAAMVRPTAAALPA
              10      20      30      40      50      60

```



```
a279.seq
  1  ATGACNCNGA  TTTGCGGCTG  CTTGATTTCA  ACGGTTTNNA  GGGCTTCGGC
51  GAGTTTGTCTG  GCGGCGGGTT  TCATGAGGCT  GCAATGGGAA  GGTACNGACA
101  CNGCGACGCG  CAGGGCGCGT  TTGGCGCCGG  CTTCTTTGGC  GGCAAGCATA
151  GCGCGCTCGA  CGGCGGCGGC  ATTGCTTGA  ATCAGCACTT  GTCCGGCGCA
201  GTTGAAGTTG  ACGGCTTCAA  CCATCTCATC  CTGTGCGGAT  TCGCGGCAAA
251  TTTGTTTTAC  CTGTTTCATC  TCCAAGCCGA  GAATCGCCGC  CATTGCGCCC
301  ACGCCTTGCG  GTACGGCGGA  CTGCATCAGT  TCGCGCGCA  NGCGCAGGAG
351  TTTGACCGCG  TCGGCAAAAT  CCAATGCGCC  GGCGGCAACN  AGTGCGGTGT
401  ATTCGCCGAN  GCTGTGTCCG  GCAACGGCGG  CAGGCGTTTT  GCCGCCGCT
451  TCCGAATAG
```

```
a279.pep
  1 MTXICGCLIS TVXRASASLS AAGFMRLQWE GTDTGSGRAR LAPASLAASI
  51 ARSTAAALPA ITTCGELKL TASTTSSCAD SAQICFTCSS SKPRIAIAIP
 101 TPCGTADCIS SARXRTSLTA SAKSNAPAAAT SAVSPXLCF ATAAGVLPPA
 151 SE*
```

	10	20	30	40	50	60
m279.pep	ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRRARLAPASLAAAMARPTAAALPA					
	:					
a279	MTXICGCLISTVXRRASASLSAAGFMRLQWEGTDTGSGRRARLAPASLAASIARSTAAALPA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m279.pep	ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA					
a279	ITTCPGELKLTASTTSSCADSAQICFTCSSSKPRIAAIAPTPCGTADCISSARXRTSLTA					
	70	80	90	100	110	120
	130	140	150			
m279.pep	SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX					
a279	SAKSNAPAATSAVYSPXLCPPATAAGVLPPASEX					
	130	140	150			

```
m519.seq (partial)
1 ..TCCGTTATCG GGCATATGGA GTTGGACAAA ACGTTTGAAG AACGCGACGA
51 AATCAACAGT ACTGTTGTTG CGGCTTTGGA CGAGGCGGCC GGGgCTTgGG
101 GTGTGAAGGT TTTGCGTTAT GAGATTAAAG ACTTGGTTCC GCCGCAAGAA
151 ATCCTTTCGT CAATGCAGGC GCAAATTACT GCCGAACGCG AAAAACGCGC
201 CCGTATCGCC GAATCCGAAG GTCGTAAATC CGAACAATC AACCTTGCCA
```

```

251 GTGGTCAGCG CGAAGCCGAA ATCCAACAAT CCGAAGGCGA GGCTCAGGCT
301 GCGGTCAATG CGTCAAATGC CGAGAAAATC GCCCGCATCA ACCGCGCCAA
351 AGGTGAAGCG GAATCCTTGC GCCTTGTGTC CGAAGCCAAT GCCGAAGCCA
401 TCCGTCAAAT TGCCGCCGCC CTTCAAACCC AAGGCGGTGC GGATGCGGTC
451 AATCTGAAGA TTGCGGAACA ATACGTCGCT GCGTTCAACA ATCTTGCCAA
501 AGAAAGCAAT ACGCTGATTA TGCCCGCCAA TGTTGCCGAC ATCGGCAGCC
551 TGATTTCTGC CGGTATGAAA ATTATCGACA GCAGCAAAAC CGCCAAaTAA

```

This corresponds to the amino acid sequence <SEQ ID 3046; ORF 519>:

```

m519.pep (partial)
1 ..SVIGRMELDK TFEERDEINS TVVAALDEAA GAWGVKVLRY EIKDLVPPQE
51 ILRSMQAQIT AEREKRARIA ESEGRKIEQI NLASGQREAE IQQSEGEAQA
101 AVNASNAEKI ARINRAKGEA ESLRLVAEAN AEAIRQIAAA LQTQGGADAV
151 NLKIAEQYVA AFNNLAKESN TLIMPANVAD IGSLSISAGMK IIDSSKTAK*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3047>:

```

g519.seq
1 atggaatttt tcattatctt gttggcagcc gtrjccgttt tcggcttcaa
51 atccttttgc gtcattcccc agcaggaagt ccacgttgtc gaaaggctcg
101 ggcgtttcca tcgcgccttg acggccggtt tgaatatctt gattcccttt
151 atcgaccgcg tcgcctaccg ccattcgctg aaagaaatcc ctttagacgt
201 acccagccag gtctgcatca cgcgcgataa tacgcaattg actggtgacg
251 gcatcatcta ttccaagta accgatccca aactcgcttc atacggttcg
301 agcaactaca ttatggcaat taccagctt gcccaaacga cgctgcgctt
351 cgttatcggg cgtatggagt tggacaaaac gtttgaagaa cgcgacgaaa
401 tcaacagtac cgtcgtctcc gccctcgatg aagccgcccg ggcttggggg
451 gtgaaagtcc tccgttacga aatcaaggat ttggttccgc cgcaagaaat
501 ccttcgcgca atgcaggcac aaattaccgc cgaacgcgaa aaacgcgccc
551 gtattgccga atccgaaggc cgtaaaatcg aacaaatcaa ccttgccagt
601 ggtcagcgtg aagccgaaat ccaacaatcc gaaggcgagg ctccagctgc
651 ggtcaatgcg tccaatgccg agaaaatcgc ccgcatacac cgcgccaaag
701 gcgaagcgga atccctgccc cttgttgccc aagccaatgc cgaagccaac
751 cgtcaaattg ccgccgccct tcaaacccaa agcggggcgg atgcggtcaa
801 tctgaagatt gcgggacaat acgttaccgc gttcaaaaat cttgccaaag
851 aagacaatac cgggattaag cccgccaggg ttgccgaaat cgggaaccct
901 aattttcggc ggcataaaaa attttcgcca gaagcaaaaa cggccaataa
951 a

```

This corresponds to the amino acid sequence <SEQ ID 3048; ORF 519.ng>:

```

g519.pep
1 MEFFIILLAA VAVFGFKSFV VIPQOEHVHV ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
201 QGREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAAN
251 RQIAAALQTQ SGADAVNLKI AGQYVTAFKN LAKEDNTRIK PAKVAEIGNP
301 NFRRHEKFSF EAKTAK*

```

ORF 519 shows 87.5% identity over a 200 aa overlap with a predicted ORF (ORF 519.ng) from *N. gonorrhoeae*:

m519/g519

```

m519.pep
10 20 30
SVIGRMELDKTFEERDEINSTVVAALDEAA
|||||
g519 YFQVTDPKLASYGSSNYIMAITQLAQTTLRVIGRMELDKTFEERDEINSTVVSALDEAA
90 100 110 120 130 140
.
40 50 60 70 80 90
m519.pep GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3049>:

This corresponds to the amino acid sequence <SEQ ID 3050; ORF 519.a>:

**m519/a519** ORFs 519 and 519.a showed a 99.5% identity in 199 aa overlap

```

                                     10      20      30
m519.pep                          SVIGRMELDKTFEERDEINSTVVAALDEAA
                                   |||||:|||||
a519      YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
           90      100      110      120      130      140

           40      50      60      70      80      90
m519.pep  GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
           |||||:|||||:|||||:|||||:|||||:|||||
a519      GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
           150      160      170      180      190      200

           100      110      120      130      140      150

```

```

m519.pep      IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a519          IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
                210      220      230      240      250      260

                160      170      180      190      200
m519.pep      NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSKTAKX
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a519          NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSKTAKX
                270      280      290      300      310

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3051>:

```

m519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGTAGCC GTCGCGGTTT TCGGTTTCAA
51  ATCCTTTGTT GTCATCCAC AACAGGAAGT CCACGTTGTC GAAAGGCTGG
101 GCGGTTTCCA TCGCGCCCTG ACGGCGGTTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC TGTGTTGCG GCTTTGGACG AGCCCGCGG GGCTTGGGGT
451 GTGAAGGTTT TGCCTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC CGAACGCGAA AACCGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTGAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCTGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAATAAA

```

This corresponds to the amino acid sequence <SEQ ID 3052; ORF 519-1>:

```

m519-1.
1  MEFFIILLVA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIFE
51  IDRVAYRHSI KEIPLDVPSQ VCITRDNTQL TVDGLIYFQV TDPKLASYGS
101 SNYIMAITQL AQTTLRVIG RMELDKTFEE RDEINSTVVA ALDEAAGAWG
151 VKVLYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3053>:

```

g519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGCAGCC GTCGCCGTTT TCGGCTTCAA
51  ATCCTTTGTC GTCATCCCCC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGGTTTCCA TCGCGCCCTG ACGGCCGTTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGATAA TACGCAATTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGATCCCA AACTCGCCTC ATACGGTTTCG
301 AGCAACTACA TTATGGCAAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG GGCTTGGGGT
451 GTGAAAGTCC TCCGTTACGA AATCAAGGAT TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCGCA ATGCAGGCAC AAATTACCGC CGAACGCGAA AACCGCGCCC
551 GTATTGCCGA ATCCGAAGGC CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTGAGCGTG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCCAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GCGAAGCGGA ATCCCTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCAA GGCGGGGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTAGCCGC GTTCAACAAT CTTGCCAAAG

```

851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG  
 901 ATTTCTGCCG GCATGA/... TATCGACAGC AGCAAAACCG CCAAATAA

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 519-1.ng>:

g519-1.pep  
 1 MEFFIILLAA VAVFGFKSFV VIPQQEVHV V ERLGRFHRAL TAGLNILIPF  
 51 IDRVA YR/SL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS  
 101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG  
 151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS  
 201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANA EAI  
 251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL  
 301 ISAGMKIIDS SKTAK\*

m519-1/g519-1 ORFs 519-1 and 519-1.ng showed a 99.0% identity in 315 aa overlap

g519-1.pep	10	20	30	40	50	60
	MEFFIILLAAVAVFGFKSFVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL					
m519-1	MEFFIILLVAVAVFGFKSFVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL					
	10	20	30	40	50	60
g519-1.pep	70	80	90	100	110	120
	KEIPLDVPSQVCITRDNTQLTVDGIIFYQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG					
m519-1	KEIPLDVPSQVCITRDNTQLTVDGIIFYQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG					
	70	80	90	100	110	120
g519-1.pep	130	140	150	160	170	180
	RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE					
m519-1	RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE					
	130	140	150	160	170	180
g519-1.pep	190	200	210	220	230	240
	KRARIAESEGRKIEQINLASGQREAEIQSEGEAQAAVNASNAEKIARINRAKGEAESLR					
m519-1	KRARIAESEGRKIEQINLASGQREAEIQSEGEAQAAVNASNAEKIARINRAKGEAESLR					
	190	200	210	220	230	240
g519-1.pep	250	260	270	280	290	300
	LVAEANA EAI RQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL					
m519-1	LVAEANA EAI RQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL					
	250	260	270	280	290	300
g519-1.pep	310					
	ISAGMKIIDSSKTAKX					
m519-1	ISAGMKIIDSSKTAKX					
	310					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3055>:

a519-1.seq  
 1 ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA  
 51 ATCCTTTGTT GTCATCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG  
 101 GCGTTTCCA TCGGCCCTG ACGGCCGTT TGAATATTTT GATTCCCTTT  
 151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT  
 201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG  
 251 GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG  
 301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC  
 351 CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA

```

401 TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCGCG AGCTTGGGGT
451 GTGAAGGTTT TCGGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAAGCGC AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCCGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 3056; ORF 519-1.a>:

```

a519-1.pep.
  1 MEFFIILLAA VVVFQKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
 51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIYFQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAELAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

**m519-1/a519-1** ORFs 519-1 and 519-1.a showed a 99.0% identity in 315 aa overlap

```

              10      20      30      40      50      60
a519-1.pep  MEFFIILLAAVVVFQKSFVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
              |||
m519-1      MEFFIILLVAVAVFQKSFVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
              10      20      30      40      50      60

              70      80      90      100     110     120
a519-1.pep  KEIPLDVPSQVCITRDNTQLTVDGIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
              |||
m519-1      KEIPLDVPSQVCITRDNTQLTVDGIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
              70      80      90      100     110     120

              130     140     150     160     170     180
a519-1.pep  RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQQAQITAERE
              |||
m519-1      RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQQAQITAERE
              130     140     150     160     170     180

              190     200     210     220     230     240
a519-1.pep  KRARIAESEGKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
              |||
m519-1      KRARIAESEGKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
              190     200     210     220     230     240

              250     260     270     280     290     300
a519-1.pep  LVAEANAELAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
              |||
m519-1      LVAEANAELAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
              250     260     270     280     290     300

              310
a519-1.pep  ISAGMKIIDSSKTAKX
              |||
m519-1      ISAGMKIIDSSKTAKX
              310

```

```
m576.seq.. (partial)
1 ..ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
51 GCAAATGAGG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTAAGCGAA
101 CCATGTCAGG AGTGATGAC GCCAAGAGAA TCAAAATGAC GCAAGAGCAG
151 GCTCAGGAAG TCATGATGAA ATTCTTTCAG GAACAACAGG CTAAGCCGT
201 AGAAAAACAC AAGCGGCAGC CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
251 TTCTGAAAG AATGCGCGC AAGACGGCG TGAAGACCAC TGCTTCGGC
301 CTGCAATACA AAATCACCAA ACAGGGCGAA GCAACACAGC GACCAAAGA
351 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTCATGAC GGTACGGTAT
401 TCGACAGCAG CAAGGCCAAG GCGCGCCCGG TCACCTTCCC TTTGAGCCAA
451 GTGATTCCGG GTTGACCGA AgCGTACAG CTCTGAAAG AAGCGCGCA
501 AGCCACGTTC TACATCCCGT CCAACCTTGC TACCGCGAA CAGGGTGC
551 GCGACAAAT CCGTCCGAAC GCCACTTGG TATTGTATGT GAAACTGGTC
601 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCACGTGCA
651 CATCAAAAAA GTAATTAA
```

```
m576.pep.. (partial)
1  ..MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
51  AQEVMKKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
101 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
151 VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDIGKPN ATLVDVFKLV
201 KIGAPENAPA KQPAOVDDIKK VN*
```

```

g576.seq. . (partial)
1  ..atgggcgttg  acatcggacg  ctccctgaaa  caaatgaagg  aacaggggcg
51  ggaatcgcgt  ttgaaagtct  ttaccgatgc  catcgaggca  gtgatcgacg
101 gcaagaagaa  caaaatgacc  gaagacgagg  ccaggagaag  gatgatgaaa
151 ttctctgcag  agcagcaggc  taaagccgta  gaaaaacaca  aggcgggatg
201 gaaggccaac  aaagaaaaaa  gcgaagcctt  cctgaaggaa  aatgcgcccg
251 aagacggcgt  gaagaccact  gcttcgggtc  tgcagtacaa  aatcaccaac
301 cagggtgaag  gcaaacagcc  gacaaaagac  gacatcgtta  ccgtggaata
351 cgaaggccgc  ctgattgcag  gtaccgtatt  cgacagcagc  aaagccaacg
401 ggcgcgccgc  cacccttcct  ttggccaag  tgattccggg  ttggaccgaa
451 ggcgctagcg  ttctgaaaga  aggcggcgaa  gcacgcttc  acatcccgtc
501 caactctgcc  taccgcgaac  agggctcggg  cgaaaaaatc  ggtcggaacg
551 ccactttgg  atttgacgtg  aaactgtgta  aaatcggcgc  acccgaaaa
601 gcgcgccgca  agcaqccgga  tcaactcgac  atcaaaaaac  taatttaa

```

g576.pep.. (partial)

1	..MGVDIGRSLK	QMKEQGAEID	LKVFTDAMQA	VYDGKEIKMT	EEQAQEVMMK
51	LQEQAQAKAV	EKKHAKADKAN	KEKGEAFLKE	NAAEDGVKTT	ASGLQYKITK
101	QGEGKQPTKD	DIVTVEYEGR	LIDGTVFDSS	KANGGPATFP	LSQVIPGWTE
151	GVRLLEKGE	ATFYIPSNLA	YREQGAGEKI	GNATLVFDV	KLKVIAPEN
201	APAKOPDOVD	IKKN*			

**m576/g576** ORFs 576 and 576.ng showed a 97.2% identity in 215 aa overlap

m576.pep                10                20                30                40                50                60  
MQQASYAMGVDIGRSLKQMKEQGAEIDLKVFTTEAMQAVYDGKEIKMTEEQAQEVMMKFLQ  
| | | | | | | | | | | | | | | | : | | | | | | | | | |

80

```

g576          MGVDIGRSLKQMKEQGAEIDLKVFDTAMQAVYDGKEIKMTEEQAQEVMMKFLQ
                10      20      30      40      50

                70      80      90      100     110     120
m576.pep      EQQAKAVEKHKADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQEGEKQPTKDDIV
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          EQQAKAVEKHKADAKANKEKGEAFLKENAAEDGVKTTASGLQYKITKQEGEKQPTKDDIV
                60      70      80      90      100     110

                130     140     150     160     170     180
m576.pep      TVEYEGRLIDGTVFDDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYRE
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          TVEYEGRLIDGTVFDDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYRE
                120     130     140     150     160     170

                190     200     210     220
m576.pep      QGAGDKIGPNATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN
                ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          QGAGEKIGPNATLVFDVKLVKIGAPENAPAKQPDQVDIKKVN
                180     190     200     210

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3061>:

```

a576.seq
1   ATGAACACCA TTTTCAAAT CAGCGCACTG ACJCTTTCCG CCGCTTTGGC
51  ACTTTCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
101 CTGCCGCCG TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
301 GCTCAGGAAG TCATGATGAA ATTCTTCAG GAACAACAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
451 CTGCAATACA AAATCACCAC ACAGGGCGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
701 GCGACAAAAA CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3062; ORF 576.a>:

```

a576.pep
1   MNTIFKISAL TLSAALALSA CGKKEAPAS ASEPAASSA QGDTSSIGST
51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
151 LQYKITQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

m576/a576 ORFs 576 and 576.a showed a 99.5% identity in 222 aa overlap

```

                10      20      30
m576.pep      MQQASYAMGV DIGRSLKQMK EQGAEIDLKV
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a576          CGKKEAPASASEPAASSAQGDTSSIGSTMQQASYAMGV DIGRSLKQMK EQGAEIDLKV
                30      40      50      60      70      80

                40      50      60      70      80      90
m576.pep      FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a576          FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                90      100     110     120     130     140

```



	100	110	120	130	140	150
m576.pep	KDGVKTTASGLQYI...KQEGEKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ					
a576	KDGVKTTASGLQYKIKQEGEKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ					
	150	160	170	180	190	200
m576.pep	VIPGWTEGVQLLKEGGEATFYIPSNLAYREQAGDKIGPNATLVFDVKLVKIGAPENAPA					
a576	VILGWTEGVQLLKEGGEATFYIPSNLAYREQAGDKIGPNATLVFDVKLVKIGAPENAPA					
	210	220	230	240	250	260
m576.pep	KQPAQVDIKKVN					
a576	KQPAQVDIKKVN					
	270					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3063>:

m576-1.seq

```

1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
51  ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
101 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTGAAAGTC TTTACCGAAG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACCGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601 GTGATTCCGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
701 GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCGGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3064; ORF 576-1>:

m576-1.pep

```

1  MNTIFKISAL TLSAALALSA CGKKEAPAS ASEPAASSA QGDTSSIGST
51  MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLEKNAE KDGVKTTASG
151 LQYKIKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3065>:

g576-1.seq

```

1  ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
51  ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
101 CTGCCGCCGC TTCTGCCGCG CAGGGCGACA CCTCTTCAAT CGGCAGCACG
151 ATGCAGCAGG CAAGCTATGC AATGGGCGTG GACATCGGAC GCTCCCTGAA
201 ACAAAATGAAG GAACAGGGCG CGGAAATCGA TTGAAAGTC TTTACCGATG
251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAATGAC CGAAGAGCAG
301 GCCCAGGAAG TGATGATGAA ATTCCTGCAG GAGCAGCAGG CTAAAGCCGT
351 AGAAAAACAC AAGGCGGATG CGAAGGCCAA CAAAGAAAAA GGCGAAGCCT
401 TCCTGAAGGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGT
451 CTGCAGTACA AAATCACCAA ACAGGGTGAA GGCAAACAGC CGACCAAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACCGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG CCACCTTCCC TTTGAGCCAA
601 GTGATTCCGG GTTGGACCGA AGGCGTACGG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG

```

701 GCGAAAAAAT CGGTCCGAAC GCCACTTTGG TATTTGACGT GAAACTGGTC  
 751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCG ATCAAGTCGA  
 801 CATCAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3066; ORF 576-1.ng>:

g576-1.pep  
 1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASAA QGDTSSIGST  
 51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTDAMQAVYD GKEIKMTEEQ  
 101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG  
 151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPATFPLSQ  
 201 VIPGWTEGVR LLKEGGEATF YIPSNLAYRE QGAGEKIGPN ATLVFDVKLV  
 251 KIGAPENAPA KQPDQVDIKK VN\*

**g576-1/m576-1** ORFs 576-1 and 576-1.ng showed a 97.8% identity in 272 aa overlap

	10	20	30	40	50	60
g576-1.pep	MNTIFKISALTLSAALALSACGKKEAAPAS	LEPAAASAAQGDTSSIGSTMQQASYAMGV				
m576-1	MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV					
	10	20	30	40	50	60
g576-1.pep	DIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH					
m576-1	DIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH					
	70	80	90	100	110	120
g576-1.pep	KADAKANKEKGEAFLENAAKDGVKTTASGLQYKITKQGEKGKQPTKDDIVTVEYEGRLID					
m576-1	KADAKANKEKGEAFLENAAKDGVKTTASGLQYKITKQGEKGKQPTKDDIVTVEYEGRLID					
	130	140	150	160	170	180
g576-1.pep	GTVFDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYREQGAGEKIGPN					
m576-1	GTVFDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYREQGAGEKIGPN					
	190	200	210	220	230	240
g576-1.pep	ATLVFDVKLVKIGAPENAPAKQPDQVDIKKVN					
m576-1	ATLVFDVKLVKIGAPENAPAKQPDQVDIKKVN					
	250	260	270			

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3067>:

a576-1.seq  
 1 ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC  
 51 ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCC GCATCT GCATCCGAAC  
 101 CTGCCGCCG TTCTTCCGCG CAGGGCGACA CCTCTCGAT CGGCAGCAGC  
 151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA  
 201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG  
 251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG  
 301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAGCCGT  
 351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT  
 401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGCG  
 451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA  
 501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT  
 551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA  
 601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGCGGA  
 651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCAG

```

701 GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3068; ORF 576-1.a>:

a576-1.pep

```

1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVTFTASG
151 LQYKITQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

**a576-1/m576-1** ORFs 576-1 and 576-1.a showed a 99.6% identity in 272 aa overlap

	10	20	30	40	50	60
a576-1.pep	MNTIFKISALT	LSAALALSAC	KGKEAAPAS	ASEPAAASSA	QGDTSSIGST	MQQASYAMGV
m576-1	MNTIFKISALT	LSAALALSAC	KGKEAAPAS	ASEPAAASSA	QGDTSSIGST	MQQASYAMGV
	10	20	30	40	50	60
	70	80	90	100	110	120
a576-1.pep	DIGRSLKQMK	EQGAEIDLKV	FTEAMQAVYD	GKEIKMTEE	QAEVMMKFL	QEQAKAVEKH
m576-1	DIGRSLKQMK	EQGAEIDLKV	FTEAMQAVYD	GKEIKMTEE	QAEVMMKFL	QEQAKAVEKH
	70	80	90	100	110	120
	130	140	150	160	170	180
a576-1.pep	KADAKANKEK	GEAFLENAA	KDGVTFTAS	GLQYKITQ	GEGKQPTK	DDIVTVEY
m576-1	KADAKANKEK	GEAFLENAA	KDGVTFTAS	GLQYKITQ	GEGKQPTK	DDIVTVEY
	130	140	150	160	170	180
	190	200	210	220	230	240
a576-1.pep	GTVFDSSKAN	GGPVTFPL	SQVILGWTE	GVQLLKEG	GEATFYIP	SNLAYREQ
m576-1	GTVFDSSKAN	GGPVTFPL	SQVILGWTE	GVQLLKEG	GEATFYIP	SNLAYREQ
	190	200	210	220	230	240
	250	260	270			
a576-1.pep	ATLVFDVKLV	KIGAPENAP	AKQPAQVD	IKKVN		
m576-1	ATLVFDVKLV	KIGAPENAP	AKQPAQVD	IKKVN		
	250	260	270			

919 gnm43.seq

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3069>:

m919.seq

```

1 ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TACGGCATCG CCGCCGCCAT
51 CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCGG CAACCCGACA
101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
151 GGAACGACGG TCGGCGGCGG CGGGGCCGTC TATACCGTTG TACCGCACCT
201 GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT
251 TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TCGCCCCAAG CCTTTCAAAC CCCCCTCCAT TCCTTTCAGG CAAAACAGTT
351 TTTTGAACGC TATTTACCGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
401 CCGGTACGGT TACCGGTAT TACGAACCGG TGCTGAAGGG CGACGACAGG

```

```

451 CGGACGGCAC AAGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCTGCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTGTGCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATACCGCCG ACCTCTCCcG ATTCCCCATC ACCGCGCGCA CAACAGCAAT
651 CAAAGGCAGG TTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCATACT CGGTTACGCC
751 GAAGACCCTG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
851 AACATCCyTA CGTTTCCATC GGACGCTATA TGGCGGATAA GGGCTACCTC
901 AAATCGGAC AAACCTCCAT GCAGGGCATT AAGTCTTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC
1051 ACGCCGCTGA TGGGGGAATA TGCCGCGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGTGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251 TGCCGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTC CTACCCAACG
1301 GTATGAAGCC CGAATACCGc CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3070; ORF 919>:

```

m919.pep
  1 MKKYLFRAL YGIAAAILAA COSKSIQTFP QPDTSVINGP DRPVGIPDPA
51 GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDGR
151 RTAQAARFPIY GIPDDFISVP LPAGLRSGKA LVRIROTGKN SGTIDNTGGT
201 HTADLSRFPI TARTTAIKGR FEGRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KSYMQRNQR LAEVLGQNPS YIFFRELGS SNGGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3071>:

```

g919.seq
  1 ATGAAAAAAC ACCTGCTCCG CTCGCCCTG TACGGcatCG CCGCCgcat
51 CctcgCCGCC TGCCAAAgca gGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCC GACCGGCCGG CCGGCATCCC CGACCCCGCC
151 GGAACGACGG TTGCCGGCGG CGGGGCCGTC TATACCGTTG TGCCGCACCT
201 GTCCATGCCC CACTGGGCGG CGCaggATTT TGCCAAAAGC CTGCAATCCT
251 TCCGCTCCG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TGCGCCCAAG CCTTTCAAAC CCCCCTGCAT TCCTTTCAGG CAAAGcGgTT
351 TTTTGAACGC TATTTACGC cgtGGCaggt tgcaggcaAC GGAAGcCTTG
401 Caggtagggt TACCGGCTAT TACGAACCGG TGCTGAAGGG CGACGGCAGG
451 CGGACGGAAC GGGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCTGCCCC CTGCCTGCCG GTTTGCGGGG CGGAAAAAAC CTGTGCCGCA
551 TCAGGCAGac ggGAAAAAAC AGCGGCACGA TCGACAATGC CGGCGGCACG
601 CATACCGCCG ACCTCTCCCG ATTCCCCATC ACCGCGCGCA CAACGGcaat
651 caaaGCGAGG TTTGAaggAA GCCGCTTCCT CCCTTACCAC ACGCGCAACC
701 AAAtcaacGG CGGCgcgcTT GACGGCAAag cccCCATCCT CggttacgcC
751 GAagaccCcg tcgaacttTT TTTTCATGCAC AtccaaggCT CGGGCCGCCT
801 GAAAACCCcg tccggcaaat acatCCGCAt cggTaagcc gacAAAAACG
851 AACAtccgTa tgtttccatc ggACGctata TGGCGGACAA AGGCTACCTC
901 AAGctcgggc agACCTCGAT GCAGGgcatc aaagcCTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC GGCAATGAGG GCCCGTCGG CGCACTGGGC
1051 ACGCCACTGA TGGGGGAATA CGCCGCGCA ATCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CAGGCAGCGC GATCAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGTTAC GGCGACGAAG CCGGCGAACT
1251 TGCCGCAAA CAGAAAACCA CGGGATACGT CTGGCAGCTC CTGCCCAACG
1301 GCATGAAGCC CGAATACCGC CCGTGA

```



```

              430      440
m919.pep      QKTTGYVWQLLPNGMKPEYRPX
              |||||
g919           QKTTGYVWQLLPNGMKPEYRPX
              430      440

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3073>:

```

a919.seq
1  ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TCGGCGCATCG CCGCCGCCAT
51 CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCGATCCC CGACCCCGCC
151 GGAACGACGG TCGGCGGCGG CGGGGCCGTT TATACCGTTG TGCCGACCTT
201 GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT
251 TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG
301 TCGGCCCAAG CCTTTCAAAC CCCCGTCCAT TCCGTTCAGG CAAAACAGTT
351 TTTTGAACGC TATTTCACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG
401 CCGGTACGGT TACCGGCTAT TACGAGCCGG GCTGAAGGG CGACGACAGG
451 CGGACGGCAC AAGCCCGCTT CCCGATTTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATACCGCCG ACCTCTCCCA ATTCCCCATC ACTGCGCGCA CAACGGCAAT
651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGGCAACCC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCATACT CGGTTACGCC
751 GAAGACCCCG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
851 AACATCCCTA CGTTTCCATC GGACGCTATA TGGCGGACAA AGGCTACCTC
901 AAGCTCGGGC AGACCTCGAT GCAGGGCATC AAAGCCTATA TGCAGCAAAA
951 CCCGCAACGC CTCGCGGAAG TTTTGGGGCA AAACCCAGC TATATCTTTT
1001 TCCGAGAGCT TACCGGAAGC AGCAATGACG GCCCTGTCCG CGCACTGGGC
1051 ACGCCGCTGA TGGGCGAGTA CGCCGGCGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGGCAGCTT CTGCCCAACG
1301 GTATGAAGCC CGAATACCGC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3074; ORF 919.a>:

```

a919.pep
1  MKKYLFRAL CGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51 GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SVQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGGDR
151 RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIQTGKN SGTIDNTGGT
201 HTADLSQFPI TARTTAIKGR FEGRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KAYMQNPQR LAEVLGQNPS YIFFRELTGS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

m919/a919 ORFs 919 and 919.a showed a 98.6% identity in 441 aa overlap

```

              10      20      30      40      50      60
m919.pep      MKKYLFRALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
              |||||
a919           MKKYLFRALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
              10      20      30      40      50      60

              70      80      90      100     110     120
m919.pep      YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
              |||||
a919           YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
              70      80      90      100     110     120

```

	130	140	150	160	170	180
m919.pep	YFTPWQVAGNGSLAGTVTGYE	PVLKGGDDRR	TAQARFPIYG	IPDDFISV	PLPAGLR	SGKA
a919	YFTPWQVAGNGSLAGTVTGYE	PVLKGGDDRR	TAQARFPIYG	IPDDFISV	PLPAGLR	SGKA
	130	140	150	160	170	180
	190	200	210	220	230	240
m919.pep	LVRIRQTGKNSGTIDNTGGT	HADLSRFP	ITARTTA	IKGRFEG	SRFLPYH	TRNQINGGAL
a919	LVRIRQTGKNSGTIDNTGGT	HADLSRFP	ITARTTA	IKGRFEG	SRFLPYH	TRNQINGGAL
	190	200	210	220	230	240
	250	260	270	280	290	300
m919.pep	DGKAPILGYAEDPVELF	FMHIQSG	RLKTPSG	KYIRIGY	ADKNEHP	YVSIGRYMADKGYL
a919	DGKAPILGYAEDPVELF	FMHIQSG	RLKTPSG	KYIRIGY	ADKNEHP	YVSIGRYMADKGYL
	250	260	270	280	290	300
	310	320	330	340	350	360
m919.pep	KLQOTSMQGIKSYMQRN	PQRLAEV	LGNPSYI	FFREL	AGSSND	GPVGALGTPLMGEYAGA
a919	KLQOTSMQGIKSYMQRN	PQRLAEV	LGNPSYI	FFREL	AGSSND	GPVGALGTPLMGEYAGA
	310	320	330	340	350	360
	370	380	390	400	410	420
m919.pep	VDRHYITLGAPLFVATA	HPVTRK	ALNRLI	MAQDTG	SAIKGAV	RVDFWGYGDEAGELAGK
a919	VDRHYITLGAPLFVATA	HPVTRK	ALNRLI	MAQDTG	SAIKGAV	RVDFWGYGDEAGELAGK
	370	380	390	400	410	420
	430	440				
m919.pep	QKTTGYVWQLLPNGMK	PEYRPX				
a919	QKTTGYVWQLLPNGMK	PEYRPX				
	430	440				

121 and 121-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3075>:

m121.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CCAATTGCTG
151 GATTTCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
401 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
451 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
501 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
551 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
601 xxxxxxCAGC TTCTTACGA CAAAACGGT GCAAAGTCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTCGCAC
701 AACGCCACCC TAAAGCAGC GGGCGGAAC TGTTTGCCAT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGGGACGCT
801 TTCCCGTTT ACCGCGCAA CCGTTGCGA CGCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGAC GCGGCATCCG CAATCCTGT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGnATTG
1001 CGTGGTTGGC GGCGTGTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA

```

This corresponds to the amino acid sequence <SEQ ID 3076; ORF 121>:

```

1 METQLYIGIM SGTSMGDGADA VLIRMDGGKW LGAEGHAFTP YPGLRRQLL
51 DLQDTGADEL HRSRLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPPL AXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
151 XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX
201 xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRTLRF TAQTVCDAYS HAAADARQMY ICDGGIRNPV
301 LTMADLAECFR TRVSLHSTAD LNLDPQWVEA AXFAWLAAAC INRIPGSPHK
351 ATGASKPCIL XAGYVY*

```

q121.seq

1	ATGGAAACAC	AGCTTTACAT	CGGCATTATG	TCGGGAACCA	GTATGGACGG
51	GGCGGATGCC	GTGCTGGTAC	GGATGGACGG	CGGCAAATGG	CTGGGCGCGG
101	AAGGGCAGCG	CTTTACCCCC	TACCTTGACC	GGTTGCGCGC	CAAATTGCTG
151	GATTTCAGG	ACACAGGCAC	AGACGAACTC	CACCCGACGA	GGATGTTGTG
201	GCAAGAATC	ACGCGCTGT	ACGCGCAAAC	CGCCGCCGAA	CTGCTGTGCA
251	GTCAAACCT	CGCTCCGTGC	GACATTACCG	CCCTCGGCTG	CCACGGGCAA
301	ACCGTCCGAC	ACGCGCCGGA	ACACGGTtac	-GTTATACGC	TTGCCGATTT
351	GCCGCTGCTG	GCGGAATGa	cgcggatttt	TACCGTCggc	gacttcCGCA
401	GCCGCGACCT	TGCTGCCGGC	GGacaAGGTG	CGCGGCTCGT	CCCGCCCTTT
451	CACGAAGCCC	TGTTCCGCGA	TGACACGGAA	ACACGCGTGG	TACTGAACAT
501	CGGCGGGATT	GCCAAACATCA	GCGTACTCCC	CCCCGGCGCA	CCCGCCTTCG
551	GCTTCGACAC	AGGGCCGGGC	AATATGCTGA	TGGAAGcgctg	gacgcaggca
601	catTGGcagc	TGCTTTACGA	CAAAAacggg	gcAAAGgcg	cacAAGGCA
651	cataTGCGc	CAACTGCTCG	cgaggctGCT	CGCCcaacCG	TATTTCTCAC
701	AACCCcaccc	aaAAGCAGC	GGgcGCGaac	TgtttgcccT	AAattggtc
751	gaaacctAcc	ttgacggcgg	cgaaaaccga	tacgacgtat	tgcggacgct
801	ttcccgattc	accgcgaaaA	ccgTttggga	cgccgctctca	CACGCAACGG
851	CAGATGCCCG	TCAAAATGAC	ATTTGCGGCG	GCGGCATCCG	CAATCTGTT
901	TTAATGGCGG	ATTTTGGGCA	ATGTTTGGCG	ACACGCGTTT	CCTTGCACAG
951	CACCGCCGAA	CTGAACCTCG	ATCCTCAATG	GGTGGAGGCG	gcgcCATTtg
1001	cgtggttggC	GGCGTGTTGG	ATTAACCGCA	TTCCCGGTAG	TCCGCACAAA
1051	GCGACCGGCG	CATCCAAACC	GTGATTCTG	GGCGCGGGAT	ATTATTATTG
1101	A				

g121.pap

1	METQLYIGIM	SGTSMGADA	VLVRMDGGKW	LGAEGHAFTF	YPDRLRRLKL
51	DIQDTGDEEL	HSRSLQSEL	SRLYAQTAAE	LLCSQNLAPC	DITALGCHGQ
101	TVRHAPHEGY	SIQLADLPQL	ANELTRIFTV	DFRSRDLAAG	QGAPLPAWFA
151	HEALFRDDRE	TRVVLNIGGI	AELISVLPGA	PAFGFDTGPG	NMLMDAVTQF
201	HWQLPYDKNG	AKAAQGNILP	QLLGRLLAHP	YFSQPHPKST	GRELFAFNWL
251	ETVLDGGENR	YDVLRLTSRF	TAQTVDVAVS	HAAADARQMY	ICGGGIRNPV
301	LMADLAEFCG	TRVSLHSTAE	LNLDPQWVEA	AAFAWLAAAC	INRIPGSPHK
351	ATGASKPCIL	GAGYYY*			

m121/g121

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDGDAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
g121	METQLYIGIMSGTSMGDGDAVLVRMDGGKWLGAEGHAFTPYPDLRRKLLDLQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPL					



```

g121      HRSRMLSQELSRLYAQTAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
           70      80      90      100      110      120
m121.pep  AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
           | : : : : :
g121      AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
           130      140      150      160      170      180
m121.pep  XXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
           : : : |||||:|||||:|||||:| ||||
g121      PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
           190      200      210      220      230      240
m121.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVASHAAADARQMYICGGGIRNPV
           |||||:|||||:|||||:|||||:|||||:| ||||
g121      GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDVASHAAADARQMYICGGGIRNPV
           250      260      270      280      290      300
           310      320      330      340      350      360
m121.pep  LMADLAECFGRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
           |||||:|||||:|||||:|||||:|||||:| ||||
g121      LMADLAECFGRVSLHSTAELNLDPOWVEAAAFWLAACWINRIPGSPHKATGASKPCIL
           310      320      330      340      350      360

m121.pep  XAGYYYY
           |||||
g121      GAGYYYY

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3079>:

```

a121.seq
1  ATGGAACAC AGCTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTGAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGCGGCCGCG GGACAAGGCG CGCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCAGACGA CCGCCCTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACCGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCC TAAAAGCACG GGGCGCGAAC TGTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGATT ACCGCGCAAA CCGTTTTCGA CGCGCTCTCA CACGCAGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGCGG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCGGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3080; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMGADA VLIRMDGGKW LGAEGHAFTP YPGLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAEE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLP LL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV

```

301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AFAWMAACW VNRIPGSPHK  
 351 ATGASKPCIL GAGYYY\*

**m121/a121** ORFs 121 and 121.a showed a 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSM	DGADAVLIRMDGGKWLGAEGHAFT	PYPGRLRRQLDLQDTGADEL			
a121	METQLYIGIMSGTSM	DGADAVLIRMDGGKWLGAEGHAFT	PYPGRLRRKLLDLQDTGADEL			
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSR	LYAQTAELLC	SQLNAPS	DIT	ALGCHGQTVR	HAP
a121	HRSRILSQELSR	LYAQTAELLC	SQLNAPS	DIT	ALGCHGQTVR	HAP
	70	80	90	100	110	120
	130	140	150	160	170	180
m121.pep	XX					
a121	AERTQIFTVGD	FRSRDLAAGGQ	GAPLVPAF	HEALFRDDRE	TRAVLNIG	GGIANISVLP
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXXXXX	QLPYDKNGAKSAQGNILPQL	LDRLLAHPYFAQRHPKST			
a121	PAFGFDTGPGNMLMDAWM	QAHWQLPYDKNGAKAAQGNILPQL	LDRLLAHPYFAQPHPKST			
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLE	TYLDGGENRYD	VLRTLSRFTAQTVCD	AVSHAAADARQMYICDGGIRNPV		
a121	GRELFAINWLE	TYLDGGENRYD	VLRTLSRFTAQTVF	DAVSHAAADARQMYICGGGIRNPV		
	250	260	270	280	290	300
	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTAD	LNLDPQWVEAAXFAWLAACW	INRIPGSPHKATGASKPCIL			
a121	LMADLAECFGTRVSLHSTAE	LNLDPQWVEAAAFWMAACW	VNRIPGSPHKATGASKPCIL			
	310	320	330	340	350	360
m121.pep	XAGYYYYX					
a121	GAGYYYYX					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3081>:

m121-1.seq

1	ATGGAAACAC	AGCTTTACAT	CGGCATCATG	TCGGAACCA	GCATGGACGG
51	GGCGGATGCC	GTA	CTGATAC	GGATGGACGG	CGGCAAATGG
101	AAGGGCAGCG	CTTTACCC	TACCCCGGCA	GGTACGCGG	CCAATTGCTG
151	GATTTGCAGG	ACACAGGCGC	AGACGAACTG	CACCGCAGCA	GGATTTTGTC
201	GCAAGAACTC	AGCCGCTAT	ATGCGCAAAC	CGCCGCCGAA	CTGCTGTGCA
251	GTCAAAACCT	CGCACCGTCC	GACATTACCG	CCCTCGGCTG	CCACGGGCAA
301	ACCGTCCGAC	ACGCGCCGGA	ACACGGTTAC	AGCATAACAGC	TTGCCGATTT
351	GCCGCTGCTG	GCGGAACGGA	CGCGGATTTT	TACCGTCGGC	GACTTCCGCA
401	GCCGCGACCT	TGCGGCCGGC	GGACAAGGCG	CGCCACTCGT	CCCCGCCTTT
451	CACGAAGCCC	TGTTCCGCGA	CAACAGGGAA	ACACGCGCGG	TACTGAACAT
501	CGGCGGGATT	GCCAACATCA	GCGTACTCCC	CCCCGACGCA	CCCGCCTTCG
551	GCTTCGACAC	AGGGCCGGGC	AATATGCTGA	TGGACGCGTG	GACGCAGGCA
601	CACTGGCAGC	TTCTTTACGA	CAAAAACGGT	GCAAAGGCGG	CACAAGGCAA
651	CATATTGCGG	CAACTGCTCG	ACAGGCTGCT	CGCCACCCG	TATTTGCGAC
701	AACCCACCC	TAAAAGCACG	GGGCGCGAAC	TGTTTGCCCT	AAATTGGCTC
751	GAAACCTACC	TTGACGGCGG	CGAAAACCGA	TACGACGTAT	TGCGGACGCT
801	TTCCCGTTTT	ACCGCGCAAA	CCGTTTGCGA	CGCCGTCTCA	CACGCAGCGG

```

mi21-1.pep
  1  METQLYIGIM  SGTSMDGADA  VLIRMDGGKW  LGAEGHAFTP  YPGLRRQLL
51  DLQDTGADEL  HRSRILSQEL  SRLYAQTAEE  LLCSQNLAPS  DITALGCHGQ
101 TVRHAPHEGY  SIQLADLPLL  AERTRIFTVP  DFRSRDLAAG  GQGAPLVPFA
151 HEALFRDNRE  TRAVLNIGGI  ANISVLPPDA  PAFGFDTGPG  NMLMDAWTQA
201 HWQLPYDKNG  AKAQAQGNILP  QLLDRLLAHP  YFAQPHPKST  GRELFALNWL
251 ETYLDGGENR  YDVLRLTSRF  TAQTVCDAVS  HAAADARQMY  ICGGGIRNFP
301 LMDALAEFCF  TRVLSHSTAD  LNLDPQWVEA  AXFAWLACW  INRIPGSPHK
351 ATGASKPCIL  XAGYYY*

```

		10	20	30	40	50	60
m121-1.pep		METQLYIGIMSGTSM	DGADAVLIRMDGGK	WLGAEGHAFTPY	PGRLRRQLLDLQ	DTGADEL	
g121							
		10	20	30	40	50	60
m121-1.pep		METQLYIGIMSGTSM	DGADAVLVRMDGGK	WLGAEGHAFTPY	PDRLLRKLDDLQ	DTGTDEL	
g121							
		70	80	90	100	110	120
m121-1.pep		HRSRILSQELSR	LYAQTAELLCSQ	NLAPSDITALGCH	GQTVRHAP	EHGYSIQLADL	PL
g121							
		70	80	90	100	110	120
m121-1.pep		HRSRILSQELSR	LYAQTAELLCSQ	NLAPCDITALGCH	GQTVRHAP	EHGYSIQLADL	PL
g121							
		130	140	150	160	170	180
m121-1.pep		AERTRIFTVGDF	SRDLAAGQGGA	PLVPAFHEALF	RDNRETRAVLN	IGGIANISVL	PPDA
g121							
		130	140	150	160	170	180
m121-1.pep		AELTRIFTVGDF	SRDLAAGQGGA	PLVPAFHEALF	RDDRETRVVLN	IGGIANISVL	PPGA
g121							
		190	200	210	220	230	240
m121-1.pep		PAFGFDTGPGN	MLMDAWTQAHW	QLPYDKNGAKA	AQGNILPQLLD	RLLAHPYFAQ	PHPKST
g121							
		190	200	210	220	230	240
m121-1.pep		PAFGFDTGPGN	MLMDAWTQAHW	QLPYDKNGAKA	AQGNILPQLL	GRLLAHPYFS	QPHPKST
g121							
		250	260	270	280	290	300
m121-1.pep		GRELFALNWLE	TYLDGGENRYD	VLRLTSRFTAQ	TVCDAVSHAA	ADARQMYICG	GGGIRNPV
g121							
		250	260	270	280	290	300
m121-1.pep		GRELFALNWLE	TYLDGGENRYD	VLRLTSRFTAQ	TVWDAVSHAA	ADARQMYICG	GGGIRNPV
g121							
		310	320	330	340	350	360
m121-1.pep		LMADLAECFGR	VSLSHSTADLN	LDPOWVEAA	FAWLAACWIN	RIPGSPHKAT	GASKPCIL
g121							
		310	320	330	340	350	360
m121-1.pep		LMADLAECFGR	VSLSHSTAE	LNLDPOWVEAA	FAWLAACWIN	RIPGSPHKAT	GASKPCIL
g121							
		310	320	330	340	350	360
m121-1.pep		XAGYYYYX					
g121							
		XAGYYYYX					

al21-1.seq  
1 ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG

```

51  GCGGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGCG GACTTCCGCA
401 GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCC GCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCC GACGCA CCGGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGT GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAAACGCT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCC TAAAGCACG GGGCGCGAAC TGTTGCCCT AAATTGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TCGGACGCT
801 TTCCCGATTG ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGACGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTA"TTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 121-1.a>:

```

a121-1.pep
1  METQLYIGIM SGTSMGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPF
151 HEALFRDDRE TRAVLNIGGI ANISVLPDA PAFGFDTPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

m121-1/a121-1 ORFs 121-1 and 121-1.a showed a 96.4% identity in 366 aa overlap

	10	20	30	40	50	60
m121-1.pep	METQLYIGIMSGTSMGADAVLIRMDGGKWLGAEGHAFTPYPGRLLRRQLLDLQDTGADEL					
a121-1	METQLYIGIMSGTSMGADAVLIRMDGGKWLGAEGHAFTPYPGRLLRRQLLDLQDTGADEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121-1.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPESYISQLADLPLL					
a121-1	HRSRMLSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPESYISVQLADLPLL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m121-1.pep	AERTQIFTVGDFRSRDLAAGGQAPLVPFHEALFRDRETRAVLNIGGIANISVLPDA					
a121-1	AERTQIFTVGDFRSRDLAAGGQAPLVPFHEALFRDDRETRAVLNIGGIANISVLPDA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121-1.pep	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLDRLLAHPYFAQPHPKST					
a121-1	PAFGFDTPGPNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121-1.pep	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICGGGIRNPV					
a121-1	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFCDVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300

```

          310      320      330      340      350      360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDLPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
          |||
a121        LMADLAECFGTRVSLHSTAE LNLDLPQWVEAAAFAWMAACWVNRI PGSPHKATGASKPCIL
          310      320      330      340      350      360

m121-1.pep  XAGYYYYX
          |||
a121        GAGYYYYX

```

## 128 and 128-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3085>:

```

m128.seq (partial)
1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACCAC
1  TACGCCAGCG AAAAAGCTGCG CGAAGCCAAA TACGCGTTC GCGAAACCGA
51  wGTCAAAAAA TAyTTCCCyG TCGGCAAwGT ATTAACCGGA CTGTTCCGCC
101 AAmTCAAAAA ACTmTACGGC ATCGGATTTA CCGAAAAAAC yGTCCCGGTC
151 TGGCACAAAG ACGTGCGCTA TtkTGAATTG CAACAAAACG GCGAAmCCAT
201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
251 CGTGGATGAA CGACTACAAA GGCGCGCGCC GTTTTTCAGA CGGCACGCTG
301 CAAyTGCCCA CCGCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGCGCG
351 CAGGGAAGCC CGCyTGAGCC ACGACGAAAT CCTCATCTC TCCACGAAA
401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
451 TCCGGCATCA ACGGCGTAKA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
501 TATGAAAAAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
551 ACGAAGAAAC CGGcgTTCCC yTGCCGAAAG AACTCTTsGA CAAAwTGCTC
601 GCCGCCAAAA ACTTCCAAsG CGGCATGTTC yTsGTCCGGC AAwTGGAGTT
651 CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
701 AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTATC
751 CAGCCGCCCG AATACAACCG CTTGCGCTTG AGCTTCGGCC ACATCTTCGC
801 AGGCGGCTAT TCCGCAGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
851 GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
901 GGCAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGnAT CGCGCAGCGG
951 nGCAGAATCC TTCAAAGCCT TCCGCGGCCG CGAACCAGC ATAGACGCAC
1001 TCTTGCGCCA CAGCGGTTT GACAAACGCG TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 3086; ORF 128>:

```

m128.pep (partial)
1  MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTGWA
51  NTVEPLTGIT ERVGRIWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH

//
1  YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGfTEKTVPV
51  WHKDVRyxEL QQNGEXIGGV YMDLYAREGK RGGAWMNDYK GRRRFSDGTL
101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
201 AAKNFQXGMF XVRQXEFALF DMMIYSEDDE GRLKNWQQVL DSVRKKVAVI
251 QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
301 GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHS GF DNAV*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3087>:

```

g128.seq
1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca
51  aatccaaacc gaagACAtca AACCCGCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCGGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGCA ACGCTTTCCC CCGCACAAA AACCAGCTC GATCACGACC
401 TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCG
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTATATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGccaaa CTGCTCGGCT TTA AAAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAACACCTCG GTCTCGCCGA CCCGCGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCCG
1101 CCAAATCAAA AAACCTTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCC
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGGCGCG
1251 CGCGTGGATG AACGACtaca AAGGCCGCCG CCGCTTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCTCAGCCC TCTTCCACGA
1401 AacCGGCCAC GGA CTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAacggcgtA GAATGGGACG CGGTGGAAT GCCCAGCCAG
1501 TTTATGGAAA ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551 CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TcgcCGCCAA AAACCTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGCTC
1701 GAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGCAAG CTATTACAGC TACGCATGGG CCGAAGTCct
1851 cAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951 gcgGCGGAAT CCTTCAAAGC CTTCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTGCTGCGC CAaagcggtT TCGACAACGC gGcttgA

```

This corresponds to the amino acid sequence <SEQ ID 3088; ORF 128.ng>:

```

g128.pep
1  MIDNALLHLG EEPFRNQIQT EDIKPAVQTA IAEARGQIAA VKAQHTTGWA
51  NTVRLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 QDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQONGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTQLQPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLPHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS ESDECLRKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSTD AYA AFESDD VAATGKRFWQ EILAVGGSRS

```

651 AAESFKAFRG REPSIDALLR QSGFDNAA\*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng)  
from *N. gonorrhoeae*:

m128/g128

	10	20	30	40	50	60
g128.pep	MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQHTGWANTVERLTGIT					
m128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQHTGWANTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
g128.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA					
m128	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130	140	150	160	170	180
g128.pep	TLSPAQKTKLDHDLRDFVLSGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128	TLSPAQKTKLNH					
	130					
	//					
				340	350	360
g128.pep	YAGEKLREAKYAFSETEVKKYFPVGKVLG					
m128	YASEKLREAKYAFSETXVKKYFPVGXVLNG					
				10	20	30
	370	380	390	400	410	420
g128.pep	LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK					
m128	LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK					
	40	50	60	70	80	90
	430	440	450	460	470	480
g128.pep	GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVDELGV					
m128	GRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVDELGV					
	100	110	120	130	140	150
	490	500	510	520	530	540
g128.pep	SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQXGMF					
m128	SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF					
	160	170	180	190	200	210
	550	560	570	580	590	600
g128.pep	LVRQMEFALFDMMIYSESDCRLKNWQQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY					
m128	XVRQXEFALFDMMIYSEDDCRLKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY					
	220	230	240	250	260	270
	610	620	630	640	650	660
g128.pep	SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS					
m128	SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS					
	280	290	300	310	320	330

```

              670      679
g128.pep      IDALLRQSGFDNAAX
              |||||:|||||:
m128          IDALLRHSGFDNAVX
              340

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3089>:

```

a128.seq
1  ATGACTGACA  ACGCACTGCT  CCATTTGGGC  GAAGAACCCC  GTTTTGATCA
51  AATCAAAACC  GAAGACATCA  AACCCGCCCT  GCAAACCGCC  ATTGCCGAAG
101 CGCGCGAACA  AATCGCGGCC  ATCAAAGCCC  AAACGCACAC  CGGCTGGGCA
151 AACACTGTCT  AACCCCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTTG
201 GGGCGTGTTG  TCGCACCTCA  ACTCCGTAC  CGACACGCC  GAACGTGCGC
251 CCGCCTACAA  TGAATTAATG  CCCGAAATTA  CCGTCTTCTT  CACCGAAATC
301 GGACAAGACA  TCGAGCTGTA  CAACCGCTTC  AAAACCATCA  AAAACTCCCC
351 CGAGTTCGAC  ACCCTCTCCC  ACGCGCAAAA  AACCAAACTC  AACCACGATC
401 TCGCGGATTT  CGTCCTCAGC  GGCGCGGAAC  TGCCGCCCGA  ACAGCAGGCA
451 GAATTGGCAA  AACTGCAAAC  CGAAGGCGCG  CAACTTTCCG  CCAAATTCTC
501 CCAAAACGTC  CTAGACGCGA  CCGACGCGTT  CGGCATTTAC  TTTGACGATG
551 CCGCACCGCT  TGCCGGCATT  CCCGAAGACG  CGCTCGCCAT  GTTTGCCGCT
601 GCGCGCAAAA  GCGAAGGCAA  AACAGGCTAC  AAAATCGGTT  TGCAGATTCC
651 GCACTACCTC  GCCGTATCC  AATACGCCGA  GCGGCAAA  CTGCGCGAAC
701 AAATCTACCG  CGCCTACGTT  ACCCGCGCCA  GCGAGCTTTC  AGACGACGGC
751 AAATTCGACA  ACACCGCCAA  CATCGACCGC  ACGCTCGAAA  ACGCCTTGCA
801 AACCGCCAAA  CTGCTCGGCT  TCAAAACTA  CGCCGAATTG  TCGCTGGCAA
851 CCAAAATGGC  GGACACCCCC  GAACAAGTTT  TAAACTTCCT  GCACGACCTC
901 GCGCGCGCG  CCAAACCTTA  CGCCGAAAAA  GACCTCGCCG  AAGTCAAAGC
951 CTTCGCGCG  GAAAGCCTCG  GCCTCGCCGA  TTTGCAACCG  TGGGACTTGG
1001 GCTACGCCG  CGAAAACTG  CGCGAAGCCA  AATACGCATT  CAGCGAAACC
1051 GAAGTCAAAA  AATACTTCCC  CGTCGGCAAA  GTATTAAACG  GACTGTTCCG
1101 CCAAATCAAA  AAATCTACG  GCATCGGATT  TACCGAAAAA  ACCGTCCCCG
1151 TCTGGCACA  AGACGTGCGC  TATTTTGAAT  TGCAACAAAA  CGGCGAAACC
1201 ATAGGCGGCG  TTTATATGGA  TTTGTACGCA  CGCGAAGGCA  AACGCGGCGG
1251 CGCGTGGATG  AACGACTACA  AAGGCGCGCG  CCGTTTTC  GACGCGACGC
1301 TGCAACTGCC  CACCGCCTAC  CTCGTCTGCA  ACTTCACCCC  GCCCGTCGGC
1351 GGCAAAGAAG  CCGCTTGAG  CCATGACGAA  ATCCTCACCC  TCTTCCACGA
1401 AACCGGACAC  GGCCTGCACC  ACCTGCTTAC  CCAAGTCGAC  GAACTGGGCG
1451 TATCCGCAT  CAACGGCGTA  GAATGGGACG  CAGTCGAACT  GCCAGTCAG
1501 TTTATGGAAA  ATTTCTGTTG  GGAATACAAT  GTCTTGCGC  AAATGTCCGC
1551 CCACGAAGAA  ACCGGCGTTC  CCCTGCCGAA  AGAACTCTTC  GACAAAATGC
1601 TCGCCGCCAA  AAATCTCAA  CGCGGAATGT  TCCTCGTCCG  CCAATGGAG
1651 TTCGCCCTCT  TTGATATGAT  GATTTACAGC  GAAGACGACG  AAGGCCGTCT
1701 GAAAAACTGG  CAACAGGTTT  TAGACAGCGT  GCGCAAAGAA  GTCGCCGTGC
1751 TCCGACCGCC  CGAATACAAC  CGCTTCGCCA  ACAGCTTCGG  CCACATCTTC
1801 GCAGGCGGCT  ATTCGCGAG  CTATTACAGC  TACGCGTGGG  CGGAAGTATT
1851 GAGCGCGGAC  GCATACGCCG  CCTTTGAAGA  AAGCGACGAT  GTCGCGGCCA
1901 CAGGCAAACG  CTTTGGCAG  GAAATCCTCG  CCGTCGGCGG  ATCGCGCAGC
1951 GCGGCAGAA  CTTCAAAGC  CTTCCGCGGA  CGCGAACCGA  GCATAGACGC
2001 ACTCTTGCGC  CACAGCGGCT  TCGACAACGC  GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3090; ORF 128.a>:

```

a128.pep
1  MTDNALLHLG  EEPFRDQIKT  EDIKPALQTA  IAEAREQIAA  IKAQHTGWA
51  NTVEPLTGIT  ERVGRIWGVV  SHLNSVTDTP  ELRAAYNELM  PEITVFFTEI
101 GQDIELYNRF  KTIKNSPEFD  TLSHAQTKL  NHDLRDFVLS  GAELPPEQQA
151 ELAKLQTEGA  QLSAKFSQNV  LDATDAFGIY  FDDAAPLAGI  PEDALAMFAA
201 AAQSEGTGY  KIGLQIPHYL  AVIQYADNRK  LREQIYRAYV  TRASELSDDG
251 KFDNTANIDR  TLENALQTAK  LLGFKNYAEL  SLATKMADTP  EQVLNLFHDL
301 ARRAKPYAEK  DLAEVKAFAR  ESLGLADLQP  WDLGYAGEKL  REAKYAFSET
351 EVKKYFPVGK  VLNLFLAQIK  KLYGIGFTEK  TVPVVHKDVR  YFELQONGET
401 IGGVYMDLYA  REGKRGGAWM  NDYKGRRRFS  DGTQLQLPTAY  LVCNFTPPVG
451 GKEARLSHDE  ILTLFHETGH  GLHHLLTQVD  ELGVSGINGV  EWDAVELPSQ
501 FMENFVWEYN  VLAQMSAHEE  TGVPLPKELF  DKMLAAKNFQ  RGMFLVRQME

```



551 FALFDMMIYS EDDEGRKKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF  
 601 AGGYSAGYYS YAWALLSAD AYAAFEESDD VAATGKRWFQ EILAVGGSRS  
 651 AAESFKAFRG REPSIDALLR HSGFDNAA\*

m128/a128 ORFs 128 and 128.a showed a 66.0% identity in 677 aa overlap

	10	20	30	40	50	60
m128.pep	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHGTGWTVEPLTGIT					
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHGTGWTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
m128.pep	ERVGRIGVGVSHLNCVADTPELRVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
a128	ERVGRIGVGVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130					
m128.pep	TLSPAQKTKLNH-----					
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
m128.pep	-----					
a128	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV					
	190	200	210	220	230	240
m128.pep	-----					
a128	TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMDTPEQVLNLFHDL					
	250	260	270	280	290	300
				140	150	
m128.pep	-----YASEKLREAKYAFSETXVKKYFPVGX					
a128	ARRAKPYAEKDLAEVKAFARESLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK					
	310	320	330	340	350	360
	160	170	180	190	200	210
m128.pep	VLNGLFAQKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM					
a128	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
	220	230	240	250	260	270
m128.pep	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHLHLLTQVD					
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHLHLLTQVD					
	430	440	450	460	470	480
	280	290	300	310	320	330
m128.pep	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ					
a128	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540
	340	350	360	370	380	390
m128.pep	XGMFXVRQXEFALFDMMIYSEDDEGRKKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	RGMFLVRQMEFALFDMMIYSEDDEGRKKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					
	550	560	570	580	590	600

m128-1.sea

This corresponds to the amino acid sequence <SEQ ID 3092; ORF 128-1>:

m128-1.pgp.

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA  
51 NTVEPLTGIT ERVGRIGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI

```

101  GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201  AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLNALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301  ARRAKPYAEK DLAEVKAFAR ESLNLADLPQ WDLGYASEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401  IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFAPPVG
451  GREARLSHDE ILILFHETGH GLHLLTQVD ELGVSGINGV EWDVAVELPSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDDEGRLLKNW QQVLDVSRKK VAVIQPPEYN RFALSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAV*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3093>:

```

g128-1.seq (partial)
1  ATGATTGACA ACGCACTGCT CCACCTTGGGC GAAGAACCCC GTTTTAAATCA
51  AATCAAAACC GAAGACATCA AACCCTGCGT CCAAACCGCC ATCGCCGAAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACCGCTCG GCAGGATTTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCTC
351 CGAATTTGCA ACGCTTTCCC CGCACAAAA AACCAAGCTC GATCAGGACC
401 TGCGCGATTT CGTATTGAGC GCGCGGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCGG CCAAAATCTC
501 CCAAACGCTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCGAAGACG CGCTCGCCAT GTTTCGCCGC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCGCG CAACCGCGAA CTGCGCGAAG
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGCG
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGCCAAA CTGCTCGGCT TTAATAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCGCGC CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAACACCTCG GTCTCGCCGA CCCGAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTTCTGGCAG GCCTGTTCCG
1101 CCAAAATCAA AACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGCGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCGC CCGCTTTGCC GACGGCACGC
1301 TGCAACTGCC AACCCTTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGCCAC GGAATGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAACGGCGTA AAA

```

This corresponds to the amino acid sequence <SEQ ID 3094; ORF 128-1.ng>:

```

g128-1.pep (partial)
1  MIDNALLHLG EEPFRNQIKT EDIKPAVQTA IAEARGQIAA VKAQHTHTGWA
51  NTVERTLGIT ERVGRWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL NHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQ WDLGYASEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV K

```

m128-1/g128-1 ORFs 128-1 and 128-1.ng showed a 94.5% identity in 491 aa overlap

10 20 30 40 50 60

100

g128-1.pep	MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGTWANTVERLTGIT
m128-1	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTWANTVEPLTGIT
	10 20 30 40 50 60
g128-1.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
m128-1	ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
	70 80 90 100 110 120
g128-1.pep	TLSPAQRTKLDHDLRDFVLSGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128-1	TLSPAQRTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
	130 140 150 160 170 180
g128-1.pep	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
	190 200 210 220 230 240
g128-1.pep	TRASELSNDGKFDNTANIDRTLLENALKTAKLLGFKNYAELSLATKMDTPEQVLNFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLANALQAKLLGFKNYAELSLATKMDTPEQVLNFLHDL
	250 260 270 280 290 300
g128-1.pep	ARRAKPYAEKDLAEVKAFAREHLGLADPQPWDLGYASEKLREAKYAFSETEVKKYFPVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLPWDLGYASEKLREAKYAFSETEVKKYFPVGK
	310 320 330 340 350 360
g128-1.pep	VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
g128-1.pep	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHGLHLLTQVD
	430 440 450 460 470 480
g128-1.pep	ELGVSGINGVK
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3095>:

a128-1.seq

1	ATGACTGACA	ACGCACTGCT	CCATTGGGC	GAAGAACCCC	GTTTGTATCA
51	AATCAAAACC	GAAGACATCA	AACCGGCCCT	GCAAACCGCC	ATTGCCGAAG
101	CGCGCGAACA	AATCGCCGCC	ATCAAAGCCC	AAACGCACAC	CGGCTGGGCA
151	AACACTGTCTG	AACCCCTGAC	CGGCATCACC	GAACGCGTCG	GCAGGATTTG
201	GGCGTGTTG	TCGCACCTCA	ACTCGGTCAC	CGACACGCC	GAACGCGCG
251	CCGCCTACAA	TGAATTAATG	CCCGAAATTA	CCGTCTTCTT	CACCGAAATC
301	GGACAAGACA	TCGAGCTGTA	CAACCGCTTC	AAAACCATCA	AAAACCTCCC
351	CGAGTTCGAC	ACCTCTCCC	ACGCGCAAAA	AACCAAATC	AACCACGATC
401	TGCGCGATTT	CGTCCTCAGC	GGCGCGGAAC	TGCCGCCCGA	ACAGCAGGCA
451	GAATTGGCAA	AACAGCAAAC	CGAAGGCGCG	CAATTTTCCG	CCAAATTCTC

```

501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCGCT TGCCTGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACC GCCAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCCG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCCG
1101 CCAAAATCAAA AAACCTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCGC
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGCGGAAACC
1201 ATAGCGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGCGCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACC GGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAAC TCCCAAGTCG
1501 TTTATGGAAT ATTTCTGTTG GGAATACAAT GTCTTGGCGC AAATGTCCCG
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAACCTTCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCCGCCCTC TTGATATGAT GATTACAGC GAAGACGACG AAGGCGGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCGGTCT
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CTTTGAAGA AAGCGACGAT GTCGCGGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAAT CCTTCAAAGC CTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3096; ORF 128-1.a>:

a128-1.pep

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVEPLTGIT ERVGRIGVVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTYG KIGLQIPHYL AVIOYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVVHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAPVLPQS
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRKLNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRWFQ EILAVGGSRS
651 AAESFKAERG REPSIDALLR HSGFDNAA*

```

**m128-1/a128-1** ORFs 128-1 and 128-1.a showed a 97.8% identity in 677 aa overlap

```

          10      20      30      40      50      60
a128-1.pep MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT
          |||
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90      100     110     120
a128-1.pep ERVGRIGVVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          |||
m128-1      ERVGRIGVVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90      100     110     120

          130     140     150     160     170     180

```

a128-1.pep	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
	130 140 150 160 170 180
a128-1.pep	190 200 210 220 230 240 FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
	190 200 210 220 230 240
a128-1.pep	250 260 270 280 290 300 TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
	250 260 270 280 290 300
a128-1.pep	310 320 330 340 350 360 ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLG YAGEKLREAKYAFSETEVKKYFPVVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLQPWDLG YASEKLREAKYAFSETEVKKYFPVVGK
	310 320 330 340 350 360
a128-1.pep	370 380 390 400 410 420 VLNGLFAQIKKLYGIGFTEKTVPVVHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVVHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
a128-1.pep	430 440 450 460 470 480 NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
	430 440 450 460 470 480
a128-1.pep	490 500 510 520 530 540 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540
a128-1.pep	550 560 570 580 590 600 RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLD SVRKEVAVVRPPEYNRFANSFGHIF
m128-1	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLD SVRKKVAVIQPPEYNRFALSFGHIF
	550 560 570 580 590 600
a128-1.pep	610 620 630 640 650 660 AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
m128-1	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
	610 620 630 640 650 660
a128-1.pep	670 679 REPSIDALLRHSGFDNAAX
m128-1	REPSIDALLRHSGFDNAVX
	670

m206.pap..

q206.seq

q206.pep

m206/q206

	10	20	30	40	50	60
m206.pep	MFP	PD	KT	FL	CL	SALL
	10	20	30	40	50	60
g206	MF	PD	KT	FL	CL	GALL
	70	80	90	100	110	120
m206.pep	LGL	IG	TP	YK	WG	GS
	70	80	90	100	110	120
g206	LGL	IG	TP	YK	WG	GS
	130	140	150	160	170	
m206.pep	LV	FF	NT	GG	AH	RY
	130	140	150	160	170	
g206	LV	FF	NT	GG	AH	RY

130 140 150 160 170

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3101>:

```
a206.seq
1  ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
51  CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101 AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151 CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
201 CTACAAATGG GCGCGCAGCA GCACCGCAAC CGGCTTCGAT TCGAGCGGCA
251 TGATTCAATT CGTTTACAAA AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301 GCCCGCGACA TGGCGGCGGC AAGCCGCAAA ATCCCCGACA GCCGCCTTAA
351 GGCCGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
401 ACGTCGGACT CTATATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451 GGCAAAACCA TCAAAACCGA AAAACTCTCC ACACCGTTTT ACGCCAAAAA
501 CTACCTCGGC GCACATACTT TCTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3102; ORF 206.a>:

```
a206.pep
1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT
101 ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 ORFs 206 and 206.a showed a 99.4% identity in 177 aa overlap

	10	20	30	40	50	60
m206.pep	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
a206	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m206.pep	LGLIGTPYKWGGSSTATGFD CSGMIQFVYK NALNVKLPRTARDMAAASRKIPDSRLKAGD					
a206	LGLIGTPYKWGGSSTATGFD CSGMIQFVYK NALNVKLPRTARDMAAASRKIPDSRLKAGD					
	70	80	90	100	110	120
	130	140	150	160	170	
m206.pep	LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX					
a206	LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3103>:

```
m287.seq
1  ATGTTTAAAC GCAGCGTAAT CGCAATGGCT TGTATTTTGT CCCTTTCAGC
51  CTGCGGGGGC GCGGTGGCG GATCGCCCGA TGTCAAGTCG GCGGACACGC
101 TGTCAAAACC TGCCGCCCTT GTTGTCTCTG AAAAAGAGAC AGAGGCAAAG
151 GAAGATGCGC CACAGGCAGG TTCTCAAGGA CAGGGCGCGC CATCCGCACA
201 AGGCAGTCAA GATATGGCGG CGGTTTCGGA AGAAAATACA GGCAATGGCG
251 GTGCGGTAAC AGCGGATAAT CCCAAAAATG AAGACGAGGT GGCACAAAAT
301 GATATGCCGC AAAATGCCGC CGGTACAGAT AGTTCGACAC CGAATCACAC
351 CCCGGATCCG AATATGCTTG CCGGAAATAT GGAAAATCAA GCAACGGATG
401 CCGGGGAATC GTCTCAGCCG GCAAACCAAC CGGATATGGC AAATGCGGCG
451 GACGGAATGC AGGGGGACGA TCCGTCGGCA GCGGGCAGAA ATGCCGGCAA
501 TACGGCTGCC CAAGGTGCAA ATCAAGCCGG AAACAATCAA GCCGCCGGTT
551 CTTCAGATCC CATCCCCGCG TCAAACCTCG CACCTGCGAA TGGCGGTAGC
```



```

601 AATTTTGGAA GGGTTGATTG GGCTAATGGC GTTTTGATTG ACGGGCCGTC
651 GCAAAATATA ACG...GACCC ACTGTAAAGG CGATTCTTGT AGTGGCAATA
701 ATTTCTTGGA TGAAGAAGTA CAGCTAAAT CAGAATTGA AAAATTAAGT
751 GATGCAGACA AAATAAGTAA TTACAAGAAA GATGGGAAGA ATGATAAATT
801 TGTCGGTTTG GTTGCCGATA GTGTGCAGAT GAAGGGAATC AATCAATATA
851 TTATCTTTTA TAAACCTAAA CCCACTTCAT TTGCGCGATT TAGGCGTCTCT
901 GCA/CGGTCGA GCGGGTCGCT TCCGGCCGAG ATGCCGCTGA TTCCCGTCAA
951 TCAGGCGGAT ACGCTGATTG TCGATGGGGA AGCGGTCAGC CTGACGGGGC
1001 ATTCGGGCAA TATCTTCGCG CCCGAAGGGA ATTACCGGTA TCTGACTTAC
1051 GGGGCGGAAA AATTGCCCGG CGGATCGTAT GCCCTTCGTG TTCAAGGCGA
1101 ACCGGCAAAA GCGGAAATGC TTGCGGGCGC GGCCGTGTAC AACGGCGAAG
1151 TACTGCATT CCATACGGAA AACGGCCGTC CGTACCCGAC CAGGGGCGAG
1201 TTTGCCGCAA AAGTCGATT CGGCAGCAA TCTGTGGACG GCATTATCGA
1251 CAGCGGCGAT GATTGCATA TGGGTACGCA AAAATTCAAA GCGGCCATCG
1301 ATGGAAACGG CTTTAAGGGG ACTTGGACGG AAAATGGCAG CGGGGATGTT
1351 TCCGGAAGT TTTACGGCCC GCGGCGGAG GAAGTGGCG GAAATACAG
1401 CTATCGCCCC ACAGATGCGG AAAAGGGCGG ATTCGGCGTG TTTGCCGGCA
1451 AAAAAGAGCA GGATTGA

```

This corresponds to the amino acid sequence <SEQ ID 3104; ORF 287>:

```

m287.pep
1 MFKRSVIAMA CIFALSACGG GGGSPDVKS ADTLSKPAAP VVSEKETEAK
51 EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT GNGGAVTADN PKNEDEVAQN
101 DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP ANQPDMANAA
151 DGMQGDPSA GGQNAAGTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
201 NFGRVDLANG VLIDGPSQNI TLTHCKGDSG SGNFLDEFV QLKSEFEKLS
251 DADKISNYKK DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS
301 ARSRRLPAE MPLIPVNQAD TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY
351 GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY NGEVLHFHTE NGRPYPTRGR
401 FAAKVDFGSK SVDGIIDSGD DLHMGTKQFK AADGNNGFKG TWTENGSGDV
451 SGKFYGPAGE EVAGKYSYRP TDAEKGFGV FAGKKEQD*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3105>:

```

g287.seq
1 atgtttaaac gcagtgtgat tgcaatggct tgtatttttc ccccttcagc
51 ctgtgggggc ggcggtggcg gatcgcccga tgtcaagtcg gcggacacgc
101 cgtcaaaacc ggccgcccc gttgttgctg aaaatgccgg ggaaggggtg
151 ctgccgaaag aaaaagaaaga tgaggaggca gcgggcggtg cgcgcgaagc
201 cgatacgagc gacgcaaccg ccggagaagg cagccaagat atggcggcag
251 ttccgcaaga aaatacaggc aatggcggtg cggcaacaac ggacaacccc
301 aaaaatgaag acgcgggggc gcaaaatgat atgccgcaaa atgccgcgca
351 atccgcaaat caaacaggga acaaccaacc cgcggttctc tcagattccg
401 ccccccgctc aaaccctgcc cctgcgaatg gcggtagcga ttttggaagg
451 acgaacgtgg gcaattctgt tgtgattgac ggaccgtcgc aaaaataaac
501 gttgacccac tgtaaaaggc attcttgtaa tggtgataat ttattggatg
551 aagaagcacc gtcaaaatca gaatttgaaa aattaagtga tgaagaaaaa
601 attaagcgat ataaaaaaga cgagcaacgg gagaattttg tccgtttggt
651 tgctgacagg gtaaaaaagg atggaactaa caaatatata atcttctata
701 cggacaaacc acctactcgt tctgcacggt cgaggaggtc gcttccggcc
751 gagattccgc tgattccggt caatcaggcc gatacgctga ttgtggatgg
801 ggaagcggtc agcctgacgg ggcattccgg caatatcttc gcgccgaag
851 ggaattaccg gtatctgact tacggggcgg aaaaattgcc cggcggatcg
901 tatgccctcc gtgtgcaagg cgaaccggca aaaggcgaaa tgcttgttgg
951 caccggcgtg tacaacggcg aagtgtctga ttcccatatg gaaaacggcc
1001 gtccgtaccc gtccggaggc aggtttgccg caaaagtcga tttcggcagc
1051 aaatctgtgg acggcattat cgacagcggc gatgatttgc atatgggtac
1101 gcaaaaattc aaagccgcga tcgatggaaa cggctttaag gggacttggg
1151 cggaaaaatg cggcggggat gtttcgggaa ggttttacgg cccggccggc
1201 gaggaagtgg cgggaaaata cagctatcgc ccgacagatg ctgaaaaggg
1251 cggattcggc gtgtttgccg gcaaaaaaga tcgggattga

```

This corresponds to the amino acid sequence <SEQ ID 3106; ORF 287.ng>:

```

g287.pep
1 MFKRSVIAMA CIFPLSACGG GGGSPDVKS ADTPSKPAAP VVAENAGEGV

```

```

51 LPKEKKDEEA AGGAPOADTQ DATAGEGSQD MAAVSAENTG NGGAATTDNP
101 KNEDAGAQND MPQNAAESAN QTGNNQPAGS SDSAPASNPA PANGGSDFGR
151 TNVGNVVID GPSQNTLTH CKGDSNCGDN LLDEEAPSKS EFEKLSDEEK
201 IKRYKKDEQR ENFVGLVADR VKKDGTKNYI IFYTDKPPTR SARSRSLPA
251 EIPLIPVQA DTLIVDGEAV SLTGHSGNIF APEGNYRYLT YGAEKLPGGS
301 YALRVQGEPA KGEMLVGTAV YNGEVLHFHM ENGRPYPSGG RFAAKVDFGS
351 KSVDDGIIDSG DDLHMGTKKF KAAIDGNGFK GTWTENGGGD VSGRFYGPAG
401 EEVAGKYSYR PTDAEKGKGGF VFAGKKDRD*

```

**m287/g287** ORFs 287 and 287.ng showed a 70.1% identity in 499 aa overlap

```

          10      20      30      40      49
m287.pep  MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPKPAAPVVSE-----KETEA
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g287      MFKRSVIAMACIFPLSACGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPEKKDEEA
          10      20      30      40      50      60

          50      60      70      80      90      100     109
m287.pep  KEDAPQAGSQGQAPSQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNMPQNAAGT
          |||  :|  |  ::|  |||||  |||||  |||||  |||||  |||||  |||||
g287      AGGAPQADTQD--ATAGEGSQDMAAVSAENTGNGGAATTDNPKNEDEVAQNMPQNAAGT
          70      80      90      100     110

          110     120     130     140     150     160     169
m287.pep  DSSTPNHTPDPNMLAGNMENQATDAGESSQPANQPDMAAADGMQGGDDPSAGGQNAGNTA
g287      -----

          170     180     190     200     210     220     229
m287.pep  AQGANQAGNNQAAGSSDPIPASNPAPANGGSNFRVLDLNGVLIDGPSQNTLTHCKGDS
          ::|||:||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g287      -ESANQTGNNQPAGSSDSAPASNPAANGGSDFGRTNVGNVVIDGPSQNTLTHCKGDS
          120     130     140     150     160     170

          230     240     250     260     270     280     289
m287.pep  CSGNNFLDEEVQLKSEFEKLSADAKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKP
          ||:|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g287      CNGDNLLEDEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKKDGTNKYIIFYTD
          180     190     200     210     220     230

          290     300     310     320     330     340     349
m287.pep  KPSTSFARFRRSARSRSLPAEMPLIPVQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLT
          ||  :  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g287      KPPT-----RSARSRSLPAEIPVQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLT
          240     250     260     270     280     290

          350     360     370     380     390     400     409
m287.pep  YGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGS
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g287      YGAEKLPGGSYALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFAAKVDFGS
          300     310     320     330     340     350

          410     420     430     440     450     460     469
m287.pep  KSVDDGIIDSGDDLHMGTKKFKAIDGNGFKGTWTENGGDVSGKFYGPAGEEVAGKYSYR
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g287      KSVDDGIIDSGDDLHMGTKKFKAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYR
          360     370     380     390     400     410

          470     480     489
m287.pep  PTDAEKGKGGFVFAGKKEQDX

```

g287  
 |||||:|  
 PTDAEKGFGVFAGKKDRDX  
 420 430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3107>:

a287.seq  
 1 ATGTTTAAAC GCAGTGTGAT TGCAATGGCT TGTATTGTTG CCCTTTCAGC  
 51 CTGTGGGGGC GCGGTGGCG GATCGCCCGA TGTAAAGTCG GCGGACACGC  
 101 TGTCAAAACC TGCCGCCCTT GTTGTACTG AAGATGTCGG GGAAGAGGTG  
 151 CTGCCGAAAG AAAAGAAAGA TGAGGAGGCG GTGAGTGGTG CGCCGCAAGC  
 201 CGATACGCAG GACGCAACCG CCGGAAAAGG CGGTCAAGAT ATGGCGGCAG  
 251 TTTCGGCAGA AAATACAGGC AATGGCGGTG CGGCAACAAC GGATAATCCC  
 301 GAAAATAAAG ACGAGGGACC GCAAAATGAT ATGCCGCAAA ATGCCGCCGA  
 351 TACAGATAGT TCGACACCGA ATCACACCCC TGCACCGAAT ATGCCAACCA  
 401 GAGATATGGG AAACCAAGCA CCGGATGCGG GGAATCGGC ACAACCGGCA  
 451 AACCAACCGG ATATGGCAAA TGCGGCGGAC GGAATGCAGG GGGACGATCC  
 501 GTCGGCAGGG GAAATGCGG GCAATACGGC AGATCAAGCT GCAATCAAG  
 551 CTGAAAACAA TCAAGTCCGC GGCTCTCAAA ATCCTGCCTC TTCAACCAAT  
 601 CCTAACGCCA CGAATGGCGG CAGCGATTTT GGAAGGATAA ATGTAGCTAA  
 651 TGGCATCAAG CTTGACAGCG GTTCGGAAAA TGTAACGTTG ACACATTGTA  
 701 AAGACAAAGT ATGCGATAGA GATTTCTTAG ATGAAGAAGC ACCACCAAAA  
 751 TCAGAAATTG AAAAATTAAG TGATGAAGAA AAAATTATAA AATATAAAAA  
 801 AGACGAGCAA CGAGAGAATT TTGTCGGTTT GGTTCGTGAC AGGGTAGAAA  
 851 AGAATGGAAC TAACAAATAT GTCATCATTT ATAAAGACAA GTCCGCTTCA  
 901 TCTTCATCTG CCGGATTCAG GCGTCTGCA CCGTCGAGGC GGTTCGCTTC  
 951 GGCCGAGATG CCGCTGATTC CCGTCAATCA GCGGATACG CTGATTGTCG  
 1001 ATGGGGAAGC GGTACGCTG ACGGGGCATT CCGGCAATAT CTTGCGCCCC  
 1051 GAAGGGAATT ACCGGTATCT GACTTACGGG GCGGAAAAAT TGTCGGCGCG  
 1101 ATCGTATGCC CTCAGTGTGC AAGGCGAACC GGCAAAAGGC GAAATGCTTG  
 1151 CGGGCACGGC CGTGTACAAC GGCGAAGTGC TGCATTTCCA TATGGAAAAC  
 1201 GGCCGTCCGT CCCCGTCCGG AGGCAGGTTT GCCGAAAAG TCGATTTCGG  
 1251 CAGCAAATCT GTGGACGGCA TTATCGACAG CGCGATGAT TTGCATATGG  
 1301 GTACGCAAAA ATTCAAAGCC GTTATCGATG GAAACGGCTT TAAGGGGACT  
 1351 TGGACGGAAG ATGGCGGCGG GGATGTTTCC GGAAGGTTT ACGCCCGGC  
 1401 CGGCGAAGAA GTGGCGGGA AATACAGCTA TCGCCGACA GATGCGGAAA  
 1451 AGGCGGATT CCGCTGTTT GCCGGCAAAA AAGAGCAGGA TTGA

This corresponds to the amino acid sequence <SEQ ID 3108; ORF 287.a>:

a287.pep  
 1 MFKRSVIAMA CIVALSAACGG GGGSPDVKS ADTLSKPAAP VVTEDVGEV  
 51 LPKEKKDEEA VSGAPQADTQ DATAGKGGQD MAAVSAENTG NGGAATDNP  
 101 ENKDEGPOND MPQNAADTDS STPNHTPAPN MPTRDMGNQA PDAGESAQPA  
 151 NQPDMANAAD GMQGDPSAG ENAGNTADQA ANQAENNQVG GSQNPASSTN  
 201 PNATNGGSDF GRINVANGIK LDSGSENVTL THCKDKVCDR DFLDEEAPPK  
 251 SEFEKLSDEE KINKYKKDEQ RENFVGLVAD RVEKNGTNKY VIIYKDKSAS  
 301 SSSARFRRSA RSRRSLPAEM PLIPVNQADT LIVDGEAVSL TGHSGNIFAP  
 351 EGNRYRLTYG AEKLSGGSYA LSVQGEPAKG EMLAGTAVYN GEVLHFHME  
 401 GRPSPSGGRF AAKVDFGSKS VDGIIIDSGDD LHMGTQKFKA VIDGNGFKGT  
 451 WTENGGGDVS GRFYGPAGEE VAGKYSYRPT DAEKGGFGVF AGKKEQD\*

m287/a287 ORFs 287 and 287.a showed a 77.2% identity in 501 aa overlap

	10	20	30	40	49
m287.pep	MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----KETEA				
	:				
a287	MFKRSVIAMACIVALSAACGGGGGGSPDVKSADTLSKPAAPVVTEDVGEVLPKEKKDEEA				
	10	20	30	40	50
	50	60	70	80	90
m287.pep	KEDAPQAGSQGQGPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDVAVQNDMPQNAAGT				
	:				
a287	VSGAPQADTQ--DATAGKGGQDMAAVSAENTGNGGAATDNPENKDEGPONDMPQNAADT				
	70	80	90	100	110

108

```

m287.pep      110      120      130      140      150      160      169
DSSTPNHTPDPNMLAGNMENQATDAGESSQPANQPDMANAADGMQGGDDPSAGGQNAGNTA
a287          110      120      130      140      150      160      169
DSSTPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMANAADGMQGGDDPSAG-ENAGNTA
120      130      140      150      160      170

m287.pep      170      180      190      200      210      220      229
AQGANQAGNNQAAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNI TLTHCKGDS
a287          170      180      190      200      210      220      229
DQAANQAENNQVGGSQNPASSTNPATNGGSDFGRLNVANGIKLDSGSENVTLTHCKDKV
180      190      200      210      220      230

m287.pep      230      240      250      260      270      280      289
CSGNNFLDEEVQLKSEFEKLSADAKISNYKKDGKNDKFVGLVADSVQMKGINQYIIIFYKP
a287          230      240      250      260      270      280      289
CD-RDFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVKNGTKNYVYIYKD
240      250      260      270      280      290

m287.pep      290      300      310      320      330      340
KP--TSFARFRRSARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRY
a287          290      300      310      320      330      340
KSASSSSARFRRSARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRY
300      310      320      330      340      350

m287.pep      350      360      370      380      390      400
LTYGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTGRFFAAKVDF
a287          350      360      370      380      390      400
LTYGAEKLSGGSYALSQGEPAKGEMLAGTAVYNGEVLHFHFMENGRPSPSGGRFAAKVDF
360      370      380      390      400      410

m287.pep      410      420      430      440      450      460
GSKSV DGIIDSGDDLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYS
a287          410      420      430      440      450      460
GSKSV DGIIDSGDDLHMGTKQKFAVIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYS
420      430      440      450      460      470

m287.pep      470      480      489
YRPTDAEKG GFGVFAGKKEQDX
a287          470      480      489
YRPTDAEKG GFGVFAGKKEQDX
480      490

```

406

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3109>:

**m406.seq**

```

1   ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG ACACTGACAG GTATTCCATC GCATGGCGGA GGTAAACGCT
101 TTGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 CACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGTACTCCA
251 TTGATGCACT GATTTCGTGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301 GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CTCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAACT GGAATATTTT GCAGTAGACA

```

```

701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACGGAAG GATTAATGGT CGATTTCTCC GATATCCGAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GTAGTGCAGC AACATAGACA
951 AGGACAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3110; ORF 406>:

**m406.pep**

```

1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSGLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPREYET AETTSAGGLTG LTSLSTLNA PALSRTQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSIGIK PTEGLMVDFS DIRPYGNHTG NSAPSVEADN
301 SHEGYGYSDE VVRQHRQGP *

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3111>:

**g406.seq**

```

1  ATGCGGGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG AACTGACAGC GTATTCCATC GCATGGCGGA GGCAAACGCT
101 TCGCGGTGCA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT GTACATTGTC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTCGCGGC GAATACATAA ACAGCCCTGC CGTCCGCACC
301 GATTACACCT ATCCGCGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACGGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA GGAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CCAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTGCA GACCGTATTT TTCTGCGCG
551 GCATAGACGT TGTCTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAACCT GGAATATTTC GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATCAAA CCGACGGAAG GATTGATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCAGC AACATAGACA
951 AGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3112; ORF 406>:

**g406.pep**

```

1  MRARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSGLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPREYET AETTSAGGLTG LTSLSTLNA PALSRTQSDG SGSRSSLGLN
151 IGGMGDYRNE TLTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSIGIK PTEGLMVDFS DIQPYGNHTG NSAPSVEADN
301 SHEGYGYSDE AVRQHRQGP *

```

ORF 406 shows 98.8% identity over a 320 aa overlap with a predicted ORF (ORF406.a) from *N. gonorrhoeae*:

**g406/m406**

	10	20	30	40	50	60
g406.pep	MRARLLIPILFSVFILSACG	TLTGIPSHGGGKRFAVEQEL	VAASARAAVK	DMDLQALHGR		
m406	MQARLLIPILFSVFILSACG	TLTGIPSHGGGKRFAVEQEL	VAASARAAVK	DMDLQALHGR		

110

	10	20	30	40	50	60
	70	80	90	100	110	120
g406.pep	KVALYIATMGDQGSGLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
m406	KVALYIATMGDQGSGLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
	130	140	150	160	170	180
g406.pep	LTTSLSTLNAPALSRTQSDGSGSRSSGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF					
m406	LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF					
	190	200	210	220	230	240
g406.pep	FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
m406	FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	250	260	270	280	290	300
g406.pep	IKPKTNAFEAAAYKENYALWMGPYKVSKGIKPTEGLMVDIFSIDIQPYGNHTGNSAPSVEADN					
m406	IKPKTNAFEAAAYKENYALWMGPYKVSKGIKPTEGLMVDIFSIRPYGNHTGNSAPSVEADN					
	310	320				
g406.pep	SHEGYGYSDEAVRQHRQGQPX					
m406	SHEGYGYSDEVVRQHRQGQPX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3113>:

```

a406.seq
1  ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG AACTGTCAG GTATTCCATC GCATGGCGGA GGTAAACGCT
101 TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301 GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACGGATGT GTTTATTAAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAACCT GGAATATTTT GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGACCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACAGAAG GATTAATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCAATATGGT AACTCTGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCAC GACATAGACA
951 AGGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3114; ORF 406.a>:

```

a406.pep
1  MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFQVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSGLTGG RYSIDALIRG EYINSPAVRT
101 DYTPRYETT AETTSGGLTG LTTSLSTLNA PALSRTQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA

```

251 AYKENYALWM GPYKVSKEGK PTEGLMVDFS DIQPYGNHMG NSAPSVEADN  
 301 SHEGYGYSDE AVRRHRQGQP \*

m406/a406 ORFs 406 and 406.a showed a 98.8% identity in 320 aa overlap

	10	20	30	40	50	60
m406.pep	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFQELVAASARAQVDMQALHGR					
a406	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFQELVAASARAQVDMQALHGR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m406.pep	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETSGGLTG					
a406	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETSGGLTG					
	70	80	90	100	110	120
	130	140	150	160	170	180
m406.pep	LTTSLSTLNAPALSRTQSDGSGSKSSLGNIIGMGDYRNETLTNPRDTAFLSHLVQTVF					
a406	LTTSLSTLNAPALSRTQSDGSGSKSSLGNIIGMGDYRNETLTNPRDTAFLSHLVQTVF					
	130	140	150	160	170	180
	190	200	210	220	230	240
m406.pep	FLRGIDVVSANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
a406	FLRGIDVVSANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m406.pep	IKPKTNAFEAAAYKENYALWMGPYKVSKEGKPTGLMVDFS DIRPYGNHMGNSAPSVEADN					
a406	IKPKTNAFEAAAYKENYALWMGPYKVSKEGKPTGLMVDFS DIQPYGNHMGNSAPSVEADN					
	250	260	270	280	290	300
	310	320				
m406.pep	SHEGYGYSDEVVRQHRQGQFX					
a406	SHEGYGYSDEAVRRHRQGQFX					
	310	320				

## EXAMPLE 2

### Expression of ORF 919

The primer described in Table 1 for ORF 919 was used to locate and clone ORF 919. The predicted gene 919 was cloned in pET vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 919-His fusion protein purification. Mice were immunized with the purified 919-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; PP, purified protein, TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm

that 919 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 919 are provided in Figure 10. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 919 and the amino acid sequence encoded thereby is provided in Example 1.

### EXAMPLE 3

#### Expression of ORF 279

The primer described in Table 1 for ORF 279 was used to locate and clone ORF 279. The predicted gene 279 was cloned in pGex vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 279-GST purification. Mice were immunized with the purified 279-GST and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 279 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 279 are provided in Figure 11. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 279 and the amino acid sequence encoded thereby is provided in Example 1.

### EXAMPLE 4

#### Expression of ORF 576 and 576-1

The primer described in Table 1 for ORF 576 was used to locate and clone ORF 576. The predicted gene 576 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 576-GST fusion protein purification. Mice were immunized with the purified 576-GST and sera



were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B).. These experiments confirm that ORF 576 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 576 are provided in Figure 12. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 576 and the amino acid sequence encoded thereby is provided in Example 1.

#### EXAMPLE 5

##### Expression of ORF 519 and 519-1

The primer described in Table 1 for ORF 519 was used to locate and clone ORF 519. The predicted gene 519 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 519-His fusion protein purification. Mice were immunized with the purified 519-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 519 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 13. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 519 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 6

Expression of ORF 121 and 121-1

The primer described in Table 1 for ORF 121 was used to locate and clone ORF 121. The predicted gene *121* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 121-His fusion protein purification. Mice were immunized with the purified 121-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 121 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 121 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 121 are provided in Figure 14. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 121 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 7

Expression of ORF 128 and 128-1

The primer described in Table 1 for ORF 128 was used to locate and clone ORF 128. The predicted gene *128* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 128-His purification. Mice were immunized with the purified 128-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D) and ELISA assay (panel E). Results show that 128 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 128 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 128 are provided in Figure 15. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J.*

*Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 128 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 8

### Expression of ORF 206

The primer described in Table 1 for ORF 206 was used to locate and clone ORF 206. The predicted gene 206 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 206-His purification. Mice were immunized with the purified 206-His and sera were used for Western blot analysis (panel B). It is worth noting that the immunoreactive band in protein extracts from meningococcus is 38 kDa instead of 17 kDa (panel A). To gain information on the nature of this antibody staining we expressed ORF 206 in *E. coli* without the His-tag and including the predicted leader peptide. Western blot analysis on total protein extracts from *E. coli* expressing this native form of the 206 protein showed a reactive band at a position of 38 kDa, as observed in meningococcus. We conclude that the 38 kDa band in panel B) is specific and that anti-206 antibodies, likely recognize a multimeric protein complex. In panel C is shown the FACS analysis, in panel D the bactericidal assay, and in panel E) the ELISA assay. Results show that 206 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 206 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 519 are provided in Figure 16. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 206 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 9

Expression of ORF 287

The primer described in Table 1 for ORF 287 was used to locate and clone ORF 287. The predicted gene 287 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 287-GST fusion protein purification. Mice were immunized with the purified 287-GST and sera were used for FACS analysis (panel B), bactericidal assay (panel C), and ELISA assay (panel D). Results show that 287 is a surface-exposed protein. Symbols: M1, molecular weight marker. Arrow indicates the position of the main recombinant protein product (A). These experiments confirm that 287 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 287 are provided in Figure 17. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 287 and the amino acid sequence encoded thereby is provided in Example 1.

## EXAMPLE 10

Expression of ORF 406

The primer described in Table 1 for ORF 406 was used to locate and clone ORF 406. The predicted gene 406 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 406-His fusion protein purification. Mice were immunized with the purified 406-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 406 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 406 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 406 are provided in Figure 18. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al.

1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 406 and the amino acid sequence encoded thereby is provided in Example 1.

### EXAMPLE 11

Table 2 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 225 among different strains.

**Table 2**

<b>225 gene variability: List of used <i>Neisseria</i> strains</b>	
<b>Identification Strains number</b>	<b>Source / reference</b>
<b>Group B</b>	
zo01_225 NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo02_225 BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zo03_225 NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo04_225 297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zo05_225 1000	R. Moxon / Seiler <i>et al.</i> , 1996
zo06_225 BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zo07_225 BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zo08_225 528	R. Moxon / Seiler <i>et al.</i> , 1996
zo09_225 NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zo10_225 BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zo11_225 NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zo12_225 NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zo13_225 NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zo14_225 NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zo15_225 SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zo16_225 NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zo17_225 NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zo18_225 BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zo19_225 BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zo20_225 44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zo21_225 MC58	R. Moxon
zo96_225 2996	Our collection
<b>Group A</b>	
zo22_225 205900	R. Moxon
zo23_225 F6124	R. Moxon
z2491 Z2491	R. Moxon / Maiden <i>et al.</i> , 1998
<b>Group C</b>	
zo24_225 90/18311	R. Moxon
zo25_225 93/4286	R. Moxon

**Others**

zo26\_225 A22 (group W) R. Moxon / Maiden *et al.*, 1998  
 zo27\_225 E26 (group X) R. Moxon / Maiden *et al.*, 1998  
 zo28\_225 860800 (group Y) R. Moxon / Maiden *et al.*, 1998  
 zo29\_225 E32 (group Z) R. Moxon / Maiden *et al.*, 1998

**Gonococcus**

zo32\_225 Ng F62 R. Moxon / Maiden *et al.*, 1998  
 zo33\_225 Ng SN4 R. Moxon  
  
 fa1090 FA1090 R. Moxon

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
 Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

>FA1090 <SEQ ID 3115>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS  
 GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
 RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDSRFLN\*

Z2491 <SEQ ID 3116>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRVPARRAGNA  
 DELIGNAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF  
 MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF  
 IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDSRFLN\*

Z001\_225 <SEQ ID 3117>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
 VKKNDSRFLN\*

Z002\_225 <SEQ ID 3118>  
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
 VKKNDSRFLN\*

Z003\_225 <SEQ ID 3119>  
 MDSFFKPAVWAVLWLMFAVRLALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
 DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
 VKKNDSRFLN\*

Z004\_225 <SEQ ID 3120>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z005\_225 <SEQ ID 3121>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z006\_225 <SEQ ID 3122>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z007\_225 <SEQ ID 3123>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z008\_225 <SEQ ID 3124>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z009\_225 <SEQ ID 3125>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z010\_225 <SEQ ID 3126>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z011\_225 <SEQ ID 3127>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF  
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF  
IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z012\_225 <SEQ ID 3128>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z013\_225 <SEQ ID 3129>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z014\_225 <SEQ ID 3130>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGFNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z015\_225 <SEQ ID 3131>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCS  
GFMQHIKFRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
RFIAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z016\_225 <SEQ ID 3132>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z017\_225 <SEQ ID 3133>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z018\_225 <SEQ ID 3134>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z019\_225 <SEQ ID 3135>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z020\_225 <SEQ ID 3136>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPINRAPARRAGNADELIGSAMGLNEQPVLPVNRVPARRAGNA  
DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF  
MQHIKFRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF  
IAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z021\_225 <SEQ ID 3137>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z022\_225 <SEQ ID 3138>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIKFRAMGINLPRTSAEQARMGTPVAR



SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z023\_225 <SEQ ID 3139>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z024\_225 <SEQ ID 3140>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z025\_225 <SEQ ID 3141>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z026\_225 <SEQ ID 3142>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z027\_225 <SEQ ID 3143>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z028\_225 <SEQ ID 3144>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z029\_225 <SEQ ID 3145>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG  
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA  
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR  
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR  
VKKNDP SRFLN\*

Z032\_225 <SEQ ID 3146>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS  
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z033\_225 <SEQ ID 3147>  
MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG  
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS  
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN  
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDP SRFLN\*

Z096\_225 <SEQ ID 3148>

```

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDPSRFLN*

```

Figure 19 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 225, further confirming its utility as an antigen for both vaccines and diagnostics.

### EXAMPLE 12

Table 3 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 235 among different strains.

**Table 3**

#### **235 gene variability: List of used *Neisseria* strains**

<b>Identification Strains number</b>	<b>Reference</b>
<b>Group B</b>	
gnmzq01 NG6/88	Seiler <i>et al.</i> , 1996
gnmzq02 BZ198	Seiler <i>et al.</i> , 1996
gnmzq03 NG3/88	Seiler <i>et al.</i> , 1996
gnmzq04 1000	Seiler <i>et al.</i> , 1996
gnmzq05 1000	Seiler <i>et al.</i> , 1996
gnmzq07 BZ169	Seiler <i>et al.</i> , 1996
gnmzq08 528	Seiler <i>et al.</i> , 1996
gnmzq09 NGP165	Seiler <i>et al.</i> , 1996
gnmzq10 BZ133	Seiler <i>et al.</i> , 1996
gnmzq11 NGE31	Seiler <i>et al.</i> , 1996
gnmzq13 NGE28	Seiler <i>et al.</i> , 1996
gnmzq14 NGH38	Seiler <i>et al.</i> , 1996
gnmzq15 SWZ107	Seiler <i>et al.</i> , 1996
gnmzq16 NGH15	Seiler <i>et al.</i> , 1996
gnmzq17 NGH36	Seiler <i>et al.</i> , 1996
gnmzq18 BZ232	Seiler <i>et al.</i> , 1996
gnmzq19 BZ83	Seiler <i>et al.</i> , 1996
gnmzq21 MC58	Virji <i>et al.</i> , 1992
<b>Group A</b>	
gnmzq22 205900	Our collection

gnmzq23	F6124	Our collection
z2491	Z2491	Maiden <i>et al.</i> , 1998

**Group C**

gnmzq24	90/18311	Our collection
gnmzq25	93/4286	Our collection

**Others**

gnmzq26	A22	(group W) Maiden <i>et al.</i> , 1998
gnmzq27	E26	(group X) Maiden <i>et al.</i> , 1998
gnmzq28	860800	(group Y) Maiden <i>et al.</i> , 1998
gnmzq29	E32	(group Z) Maiden <i>et al.</i> , 1998
gnmizq31	<i>N. lactamica</i>	Our collection

**Gonococcus**

gnmzq32	Ng F62	Maiden <i>et al.</i> , 1998
gnmzq33	Ng SN4	Our collection
fa1090	FA1090	Dempsey <i>et al.</i> 1991

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.  
Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279  
Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3149>  
MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVGAVVNQIANSLT  
DRGYQVSKTAAAYNLLSPYSRNGILKGPRFVEEQPK\*

GNMZQ01 <SEQ ID 3150>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ02 <SEQ ID 3151>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ03 <SEQ ID 3152>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST

AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ04 <SEQ ID 3153>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ05 <SEQ ID 3154>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ07 <SEQ ID 3155>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ08 <SEQ ID 3156>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ09 <SEQ ID 3157>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ10 <SEQ ID 3158>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ11 <SEQ ID 3159>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ13 <SEQ ID 3160>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ14 <SEQ ID 3161>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ15 <SEQ ID 3162>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ16 <SEQ ID 3163>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ17 <SEQ ID 3164>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ18 <SEQ ID 3165>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ19 <SEQ ID 3166>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ21 <SEQ ID 3166>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ22 <SEQ ID 3167>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ23 <SEQ ID 3168>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ24 <SEQ ID 3169>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ25 <SEQ ID 3170>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ26 <SEQ ID 3171>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ27 <SEQ ID 3172>  
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS  
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ28 <SEQ ID 3173>  
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ29 <SEQ ID 3174>  
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ31 <SEQ ID 3175>  
 MKPLILGLAAVLALSACQVQKAPDFDYTAFKESKPASILVVPPLNESPDVNGTWGMLAST  
 AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITITEYGT  
 YQILDSVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
 DRGYQVSKTAAAYDLLSPYSHNGILKGPRFVEEQPK\*

GNMZQ32 <SEQ ID 3176>  
 MKPLILGLAAVLALSACQVRKAPDLDTYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
 AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK\*

GNMZQ33 <SEQ ID 3177>  
 MKPLILGLAAVLALSACQVRKAPDLDTYTSFKESKPASILVVPPLNESPDVNGTWGMLAST  
 AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK\*

Z2491 <SEQ ID 3178>  
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST  
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT  
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT  
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK\*

Figure 20 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 235, further confirming its utility as an antigen for both vaccines and diagnostics.

### EXAMPLE 13

Table 4 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 287 among different strains.

**Table 4**

#### **287 gene variability: List of used *Neisseria* strains**

<b>Identification Strains number</b>	<b>Reference</b>
--	------------------

**Group B**

287_2	BZ198	Seiler <i>et al.</i> , 1996
287_9	NGP165	Seiler <i>et al.</i> , 1996
287_14	NGH38	Seiler <i>et al.</i> , 1996
287_21	MC58	Virji <i>et al.</i> , 1992

**Group A**

z2491	Z2491	Maiden <i>et al.</i> , 1998
-------	-------	-----------------------------

**Gonococcus**

fa1090	FA1090	Dempsey <i>et al.</i> 1991
--------	--------	----------------------------

**References:**

- Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.  
Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279  
Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

287\_14 <SEQ ID 3179>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
QGAPSAQGGQDMAAVSEENTGNNGAAATDKPKNEDEGAQNDMPQNAADTDSLTPNHTPAS  
NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNQ  
TAGSQNPASSTNPSATNSGGDFGRNTVGNSSVIDGPSQNTLTHCKGDCSGNNFLDEEV  
QLKSEFEKLSADAKISNYKKDGKNDKDFVGLVADSVQMKGINQYIIIFYKPKPTSFAR  
FRRSARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAELP  
GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSRGRFAAKVDFGSKSVDGII  
DSGDGLHMGTOKEKAAIDGNGFKGTWTENGGDVSGKFYGPAGEEVAGKYSYRPTDAEKG  
GFGVFAGKKEQD\*

287\_2 <SEQ ID 3180>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
QGAPSAQGGQDMAAVSEENTGNNGAAATDKPKNEDEGAQNDMPQNAADTDSLTPNHTPAS  
NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNQ  
TAGSQNPASSTNPSATNSGGDFGRNTVGNSSVIDGPSQNTLTHCKGDCSGNNFLDEEV  
QLKSEFEKLSADAKISNYKKDGKNDKDFVGLVADSVQMKGINQYIIIFYKPKPTSFAR  
FRRSARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAELP  
GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSRGRFAAKVDFGSKSVDGII  
DSGDGLHMGTOKEKAAIDGNGFKGTWTENGGDVSGKFYGPAGEEVAGKYSYRPTDAEKG  
GFGVFAGKKEQD\*

287\_21. <SEQ ID 3181>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG  
QGAPSAQGSQDMAAVSEENTGNNGAVTADNPKNEDEVAQNDMPQNAAGTDSSTPNHTPDP  
NMLAGNMENQATDAGESSQPANQPDMAANTADGMQGGDDPSAGGQAGNTAAQGANQAGNNQ  
AAGSSDPIPASNPAPANGGSNFGFVLDLGVLDGPSQNTLTHCKGDCSGNNFLDEEV  
QLKSEFEKLSADAKISNYKKDGKNDKDFVGLVADSVQMKGINQYIIIFYKPKPTSFARFRRS

ARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNRYRLTYGAEKLPGGSY  
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVGGIIDSGD  
DLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGFGV  
FAGKKEQD\*

287\_9 <SEQ ID 3182>  
MFKRSVIAMACIVALSACGGGGGSPDVKSADTLSPKPAAPVVTEDVGEVLPKEKKDEEA  
VSGAPQADTQDATAGKGGQDMAAVSAENTGNGGAATTDNPENKDEGPQNDMPQNAADTDS  
STPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMAADGMQGGDDPSAGENAGNTADQA  
ANQAENNVQVGGSQNPASSTNPNATNGGSDFGFRINVANGIKLDSGSENVTLTHCKDKVCDR  
DFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVKNGTNGKYVITYKDKSAS  
SSSARFRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNRYRLTYG  
AEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKS  
VDGIIDSGDDLHMGTKQKFAVIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPT  
DAEKGFGVFAGKKEQD\*

FA1090 <SEQ ID 3183>  
MFKRSVIAMACIFPLSACGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPEKKDEEA  
AGGAPQADTQDATAGEGSDMAAVSAENTGNGGAATTDNPKNEDEGAQNMPQNAAESAN  
QTGNNQFAGSSDSAPASNPAFANGGSDFGRTNVGNSVVIDGSPQNTLTHCKGDSNGDN  
LLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKNGTNGKYVITYKDKPPT  
SARSRRSLPAEIPVQADTLIVDGEAVSLTGHSNIFAPEGNRYRLTYGAEKLPGGS  
YALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFA...VFGSKSVGGIIDSG  
DDLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPTDAEKGFGV  
VFAGKKDRD\*

Z2491 <SEQ ID 3184>  
MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPKPAAPVSEKETEAKEDAPQAGSQG  
QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNMPQNAAGTDSSTPNHTPDP  
NMLAGNMENQATDAGESSQPANQPDMAADGMQGGDDPSAGGQNAAGTAAQGANQAGNNQ  
AAGSSDPIPASNPAPANGGSDFGFVLDLNGVLIDGSPQNTLTHCKGDSNGNFDLDEEV  
QLKSEFEKLSADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFARFRS  
ARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNRYRLTYGAEKLPGGSY  
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVGGIIDSGD  
DLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGFGV  
FAGKKEQD\*

Figure 21 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 287, further confirming its utility as an antigen for both vaccines and diagnostics.

#### EXAMPLE 14

Table 5 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 519 among different strains.

**Table 5**

#### **519 gene variability: List of used *Neisseria* strains**

Identification Strains	Source / reference
------------------------	--------------------



**number****Group B**

zv01_519	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv02_519	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zv03_519ass	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv04_519	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zv05_519	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zv06_519ass	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zv07_519	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zv11_519	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zv12_519	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zv18_519	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zv19_519	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zv20_519ass	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zv21_519ass	MC58	R. Moxon
zv96_519	2996	Our collection

**Group A**

zv22_519ass	205900	R. Moxon
z2491_519	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

**Others**

zv26_519	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zv27_519	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zv28_519	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zv29_519ass	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998

**Gonococcus**

zv32_519	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
fa1090_519	FA1090	R. Moxon

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090\_519 <SEQ ID 3185>  
MEFFIILLAAVAVFGFKSFVVIQQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGAEESLR  
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

Z2491\_519 &lt;SEQ ID 3186&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV01\_519 &lt;SEQ ID 3187&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV02\_519 &lt;SEQ ID 3188&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV03\_519 &lt;SEQ ID 3189&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV04\_519 &lt;SEQ ID 3190&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV05\_519 &lt;SEQ ID 3191&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV06\_519ASS &lt;SEQ ID 3192&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV07\_519 &lt;SEQ ID 3193&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV11\_519 &lt;SEQ ID 3194&gt;

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV12\_519 &lt;SEQ ID 3195&gt;

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV18\_519 &lt;SEQ ID 3196&gt;

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV19\_519 &lt;SEQ ID 3197&gt;

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV20\_519ASS &lt;SEQ ID 3198&gt;

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSM  
ISAGMKIIDSSKTAK\*

ZV21\_519ASS &lt;SEQ ID 3199&gt;

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV22\_519ASS &lt;SEQ ID 3200&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV26\_519 &lt;SEQ ID 3201&gt;

MEFFIILLVAVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL  
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG  
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE  
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR  
LVAEANAFAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL  
ISAGMKIIDSSKTAK\*

ZV27\_519 &lt;SEQ ID 3202&gt;

```
MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGI IYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGKIEQINLASGQREAEIQQSEGEQAQAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

ZV28\_519 <SEQ ID 3203>

```
MEFFIILLAAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGI IYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGKIEQINLASGQREAEIQQSEGEQAQAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

ZV29\_519ASS <SEQ ID 3204>

```
MEFFIILLAAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGI IYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGKIEQINLASGQREPEIQQSEGEQAQAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSNKTAK*
```

ZV32\_519 <SEQ ID 3205>

```
MEFFIILLAAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGI IYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGKIEQINLASGQREAEIQQSEGEQAQAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

ZV96\_519 <SEQ ID 3206>

```
MEFFIILLAAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGI IYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGKIEQINLASGQREAEIQQSEGEQAQAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*
```

Figure 22 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 519, further confirming its utility as an antigen for both vaccines and diagnostics.

#### EXAMPLE 15

Table 6 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 919 among different strains.

**Table 6**

#### **919 gene variability: List of used *Neisseria* strains**

Identification Strains	Source / reference
------------------------	--------------------

**number****Group B**

zm01	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm02	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zm03	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm04	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zm05	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zm06	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zm07	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zm08n	528	R. Moxon / Seiler <i>et al.</i> , 1996
zm09	NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zm10	BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zm11asbc	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zm12	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zm13	NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zm14	NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zm15	SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zm16	NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zm17	NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zm18	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zm19	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zm20	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zm21	MC58	R. Moxon
zm96	2996	Our collection

**Group A**

zm22	205900	R. Moxon
zm23asbc	F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

**Group C**

zm24	90/18311	R. Moxon
zm25	93/4286	R. Moxon

**Others**

zm26	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zm27bc	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zm28	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zm29asbc	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998
zm31asbc	<i>N. lactamica</i>	R. Moxon

**Gonococcus**

zm32asbc	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
zm33asbc	Ng SN4	R. Moxon

fa1090	FA1090	R. Moxon
--------	--------	----------

**References:**

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.  
 Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3207>  
 MKKHLRSALYGIAAAIILAACQSRSIQTFFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV  
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTERARFPIYGI PDDFI SVPLPAGLRGGKN  
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNLIPYVSI GRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGFPVGALGTPLMGEYAGA  
 IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

Z2491 <SEQ ID 3208>  
 MKKYLFRALCGIAAAIILAACQSKSIQTFFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTAQARFPIYGI PDDFI SVPLPAGLRSGKA  
 LVRI RQTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
 KLGQTSMQGIKAYMQNPQRLAEVLGQNPSYIFFREL TGSNDGFPVGALGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM01 <SEQ ID 3209>  
 MKKYLFRALYGIAAAIILAACQSKSIQTFFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTAQARFPIYGI PDDFI SVPLPAGLRSGKA  
 LVRI RQTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGFPVGALGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM02 <SEQ ID 3210>  
 MKKYLFRALYGIAAAIILAACQSKSIQTFFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTAQARFPIYGI PDDFI SVPLPAGLRSGKA  
 LVRI RQTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGFPVGALGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM03 <SEQ ID 3211>  
 MKKYLFRALYGIAAAIILAACQSKSIQTFFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEPEVLKGDGRRTAQARFPIYGI PDDFI SVPLPAGLRSGKA  
 LVRI RQTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSI GRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGFPVGALGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM04 <SEQ ID 3212>

MKKYLFRAALCGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMQQNQRRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM05 <SEQ ID 3213>

MKKYLFRAALYGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM06 <SEQ ID 3214>

MKKYLFRAALYGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM07 <SEQ ID 3215>

MKKYLFRAALYGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM08N <SEQ ID 3216>

MKKYLFRAALYGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM09 <SEQ ID 3217>

MKKYLFRAALCGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRRLAEVLGQNPSYIFFRELTSNDGPGVAGLGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM10 <SEQ ID 3218>

MKKYLFRAALCGIAAAIILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYYPEVLKGGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA  
 LVRIQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL

KLGTSMQGIKSYMRONPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM11ASBC <SEQ ID 3219>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLGTSMQGIKSYMRONPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM12 <SEQ ID 3220>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGTSMQGIKSYMRONPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM13 <SEQ ID 3221>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAEQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
KLGTSMQGIKAYMRONPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM14 <SEQ ID 3222>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLGTSMQGIKSYMRONPQRLAEVLGQNPSYIFFRELTGSRNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM15 <SEQ ID 3223>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDLAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLGTSMQGIKSYMRONPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM16 <SEQ ID 3224>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPGRPVGIPDPAGTTVGGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIRQTGKNSGTIDNTGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  
KLGTSMQGIKSYMRONPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM17 <SEQ ID 3225>



MKKYLFRAALYGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM18 <SEQ ID 3226>

MKKYLFRAALYGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM19 <SEQ ID 3227>

MKKYLFRAALYGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM20 <SEQ ID 3228>

MKKYLFRAALYGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM21 <SEQ ID 3229>

MKKYLFRAALYGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM22 <SEQ ID 3230>

MKKYLFRAALCGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMQNPQRLAEVLGQNPSYIFFRELTSNDGPGVAGLTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM23ASBC <SEQ ID 3231>

MKKYLFRAALYGIAAAILAACQSKSIQTFFPQDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL

KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGETAGK  
MKEPGYVWQLLPNGMKPEYRP\*

ZM24 <SEQ ID 3232>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM25 <SEQ ID 3233>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIKYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM26 <SEQ ID 3234>

MKKYLFRAALYGISAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIKRYMADKGYL  
KLQOTSMQGIKAYMQQNQRLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM27BC <SEQ ID 3235>

MKKYLFRAALYGISAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIQRTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIKRYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGETAGK  
MKEPGYVWQLLPNGMKPEYRP\*

ZM28 <SEQ ID 3236>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIKRYMADKGYL  
KLQOTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM29ASBC <SEQ ID 3237>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVAGGGAV  
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER  
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA  
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
DGKAPILGYAEDPVELFFMHIIQSGRLKTPSGKYIRIGYADKNEHPYVSIKRYMADKGYL  
KLQOTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA  
VDRHYITLGAPLFVATTHPIRTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
QKTTGYVWQLLPNGMKPEYRP\*

ZM31ASBC <SEQ ID 3238>

MKKHLFRAALYGIAAAILAACQSKSIQTFFPQPDTSIIKGPDRPAGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKOFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYVFFRELAGSGNDGFPVGALGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM32ASBC <SEQ ID 3239>

MKKHLLRSALYGIAAAILAACQSRSIQTFFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV  
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTARARFPIYGI PDDFISVPLPAGLRGGKA  
 LVRIQRTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYI FFRELAGSGGDPVGALGTPLMGGYAGA  
 IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM33ASBC <SEQ ID 3240>

MKKHLLRSALYGIAAAILAACQSRSIQTFFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV  
 YTVVPHLSMPHWAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTARARFPIYGI PDDFISVPLPAGLRGGKN  
 LVRIQRTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYI FFRELAGSGNEGDPVGALGTPLMGEYAGA  
 IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

ZM96 <SEQ ID 3241>

MKKYLFRAALYGIAAAILAACQSKSIQTFFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV  
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKOFFER  
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA  
 LVRIQRTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL  
 DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL  
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYI FFRELAGSSNDGFPVGALGTPLMGEYAGA  
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK  
 QKTTGYVWQLLPNGMKPEYRP\*

Figure 23 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 919, further confirming its utility as an antigen for both vaccines and diagnostics.

#### EXAMPLE 16

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

**Table 7:** Oligonucleotides used for PCR to amplify complete or partial ORFs

ORF	primer	Sequence	Restriction sites
001	Forward	CGCGGATCCCATATG-TGGATGGTGCTGGTCAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCCGTCTTGTCAC	XhoI
003	Forward	CGCGGATCCCATATG-GTCGTATTCGTGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAAATCATGAACACGCGC	XhoI
005	Forward	CGCGGATCCCATATG-GACAATATTGACATGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATCACATCCGCCCCG	XhoI
006	Forward	CGCGGATCCCATATG-CTGCTGGTGCTGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGTTCCGGCTTTGATGT	XhoI
007	Forward	CGCGGATCCCATATG-GCCGACAACAGCATCAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAGGCGTTCATGATATAAG	XhoI
008	Forward	CGCGGATCCCATATG-AACAACAGACATTTTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCTGTCCGGTAAAAGAC	XhoI
009	Forward	CGCGGATCCCATATG-CCCCGCGCTGCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGGCTTTTGCCACGTTTT	XhoI
011	Forward	CGCGGATCCCATATG-AAGACACACCGCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGGTCAGTACGGT	XhoI
012	Forward	CGCGGATCCCATATG-CTCGCCCGTTGCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCGGGGAAGAGGCAC	XhoI
013	Forward	CGCGGATCCCATATG-CCTTTGACCATGCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGATTCGGCAAAAAAATCT	XhoI
018	Forward	CGCGGATCCCATATG-CAGCAGAGGCAGTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGAGGCGAACGCC	XhoI
019	Forward	AAAGAATTC-CTGCCAGCCGGCAAGACCCCGGC	Eco RI
	Reverse	AAACTGCAG-TCAGCGGGCGGGGACAATGCCCAT	Pst I
023	Forward	AAAGAATTC-AAAGAATATTCGGCATGGCAGGC	Eco RI
	Reverse	AAACTGCAG-TTACCCCCAAATCACTTTAACTGA	Pst I
025	Forward	AAAGAATTC-TGCGCCACCCAACAGCCTGCTCC	Eco RI
	Reverse	AAACTGCAG-TCAGAACGCGATATAGCTGTTCGG	Pst I
031	Forward	CGCGGATCCCATATG-GTCTCCCTTCGCTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATGTAAGACGGGGACAAC	XhoI
032	Forward	CGCGGATCCCATATG-CGGCGAAACGTGC	BamHI-

	Reverse	CCCGCTCGAG-CTGGTTTTTTTGATATTTGTG	NdeI
033	Forward	CGCGGATCCCATATG-GCGGCGGCAGACA	XhoI
			BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ATTTGCCGCATCCCGAT	XhoI
034	Forward	CGCGGATCCCATATG-GCCGAAAACAGCTACGG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGACGATTTGGTTCAATT	XhoI
036	Forward	CGCGGATCCCATATG-CTGAAGCCGTGCG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CCGGAATGCGTATCGG	XhoI
038	Forward	CGCGGATCCCATATG-ACCGATTTCCGCCA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTCTACGCCGTACTGCC	XhoI
039	Forward	CGCGGATCCCATATG-CCGTCCGAACCGC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TAGGATGACGAGGTAGG	XhoI
041	Forward	CGCGGATCCCATATG-TTCGTGCGGAACCGC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GCCCAAAAACCTTTTCAA	XhoI
042	Forward	CGCGGATCCCATATG-ACGATGATTTGCTTGC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGCAGCCTGCATTTGAC	XhoI
043	Forward	AAAAAAGGTACC-ATGGTTGTTCAAATCAAATATC	Kpn I
	Reverse	AAACTGCAG-TTATTGCGCTTCACCTTCCGCCGC	Pst I
043a	Forward	AAAAAAGGTACC-GCAAAAGTGATGGCGGCTTGGACGGTGC	Kpn I
	Reverse	AAAAAACTGCAG-	Pst I
		TTAATCCTGCAACACGAATTCGCCCCGTCCG	
044	Forward	CGCGGATCCCATATG-CCGTCCGACTAGAG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ATGCGCTACGGTAGCCA	XhoI
046	Forward	AAAGAATTC-ATGTCGGCAATGCTCCCGACAAG	Eco RI
	Reverse	AAACTGCAG-TCACTCGCGACCCACACCGTGAA	Pst I
047	Forward	CGCGGATCCCATATG-GTCATCATAACAGGCG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCCGAAAAAGCCCATTTTG	XhoI
048	Forward	AAAGAATTC-ATGCTCAACAAAGGCGAAGAATTGCC	Eco RI
	Reverse	AAACTGCAG-TCAAGATTCGACGGGGATGATGCC	Pst I
049	Forward	AAAGAATTC-ATGCGGGCGCAGGCGTTTGATCAGCC	Eco RI
	Reverse	AAACTGCAG-AAGGCGTATCTGAAAAAATGGCAG	Pst I
050	Forward	CGCGGATCCCATATG-GGCGCGGGCTGG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AATCGGGCCATCTTCGA	XhoI
052	Forward	AAAAAAGAATTC-ATGGCTTTGGTGGCGGAGGAAAC	Eco RI
	Reverse	AAAAAAGTCGAC-TCAGGCGGCGTTTTTCACCTTCCT	Sal I
052a	Forward	AAAAAAGAATTC-GTGGCGGAGGAAACGGAAATATCCGC	Eco RI

	Reverse	AAAAAACTGCAG-TTAGCTGTTTTTGGAAACGCCGTCCAACCC	Pst I
073	Forward	CGCGGATCCCATATG-TGTATGCCATATAAGAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACCGGATTGTCCGAC	XhoI
075	Forward	CGCGGATCCCATATG-CCGTCTTACTTCATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCACCAATGCCGATTATTT	XhoI
077a	Forward	AAAAAAGAATTC-GGCGGCATTTTCATCGACACCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGAACATCTGCACAAACGCAAT	Pst I
080	Forward	AAAGAATTC-GCGTCCGGGCTGGTTTGGTTTACAATTC	Eco RI
	Reverse	AAACTGCAG-CTATTCCTCGGATTCTTTTTCGGG	Pst I
081	Forward	AAAGAATTC-ATGAAACCACTGGACCTAAATTTTCATCTG	Eco RI
	Reverse	AAACTGCAG-TCACTTATCCTCCAATGCCTC	Pst I
082	Forward	AAAGAATTC-ATGTGGTTGTTGAAGTGCCTGC	Eco RI
	Reverse	AAACTGCAG-TTACGCGGATTTCGGCAGTTGG	Pst I
084	Forward	AAAGAATTC-TATCACCCAGAATATGAATACGGCTACCG	Eco RI
	Reverse	AAACTGCAG-TTATACTTGGGCGCAACATGA	Pst I
085	Forward	CGCGGATCCCATATG-GGTAAAGGGCAGGACT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CAAAGCCTTAAACGCTTCG	XhoI
086	Forward	AAAAAAGGTACC-TATTTGGCATCAAAGAAGGCGG	Kpn I
	Reverse	AAACTGCAG-TTACTCCACCCGATAACCGCG	Pst I
087	Forward	AAAGAATTC-ATGGGCGGTAAAACCTTTATGC	Eco RI
	Reverse	AAACTGCAG-TTACGCCGCACACGCAATCGC	Pst I
087a	Forward	AAAAAAGAATTC-AAGCTATTAGGCGTGCCGATTGTGATTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCTGCAAGATGCCAGCTTGCC	Pst I
088	Forward	AAAAAAGAATTC-ATGTTTTTATGGCTCGCACATTTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAGCGGATTTTGAGGGTACTCAAACC	Pst I
089	Forward	CGCGGATCCCATATG-CCGCCAAAATCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCGCATACCAAAGCCA	XhoI
090	Forward	CGCGGATCCCATATG-CGCATAGTCGAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCAAAACGGCGGTACG	XhoI
091	Forward	AAAGAATTC-ATGGAAATACCCGTACCGCCGAGTCC	Eco RI
	Reverse	AAACTGCAG-TCAGCGCAGGGGGTAGCCCAAGCC	Pst I
092	Forward	AAAGAATTC-ATGTTTTTTATTTCAATCCG	Eco RI
	Reverse	AAACTGCAG-TCAAATCTGTTTCGACAATGC	Pst I
093	Forward	AAAGAATTC-ATGCAGAATTTTGGCAAAGTGGC	Eco RI
	Reverse	AAACTGCAG-CTATGGCTCGTCATACCGGGC	Pst I
094	Forward	AAAGAATTC-ATGCCGTCACGGAAGCGCATCAACTC	Eco RI
	Reverse	AAACTGCAG-TTATCCCGGCCATACCGCCGAACA	Pst I
095	Forward	AAAGAATTC-ATGTCCTTTTCATTTGAACATGGACGG	Eco RI
	Reverse	AAACTGCAG-TCAACGCCGCAGGCACTAACGCCC	Pst I
096	Forward	AAAGAATTC-ATGGCTCGTCATACCGGGCAGGG	Eco RI

	Reverse	AAACTGCAG-TCAAAGGAAAAGGCCGTCTGAAAAGCG	Pst I
097	Forward	AAAGAATTC-ATGGACACTTCAAAACAAACACTGTTG	Eco RI
	Reverse	AAACTGCAG-TCAGCCCAAATACCAGAATTTTCAG	Pst I
098	Forward	AAAGAATTC-GATGAACGCAGCCAGCATGGATACG	Eco RI
	Reverse	AAACTGCAG-TTACGACATTCTGATTTGGCA	Pst I
102	Forward	AAAAAAGAATTC-GGCCTGATGATTTTGGAAAGTCAACAC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
105	Forward	CGCGGATCCCATATG-TCCGCAAACGAATACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGTCTGCCAGTTTCAG	XhoI
107	Forward	AAAAAAGAATTC- CTGATGATTTTGGAAAGTCAACACCCATTATCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
107b	Forward	AAAAAAGAATTC- GATACCCAAGCCCCCGCCGGCACAAACTACTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCGTCGCCTTTAAAGTATTTGAGCAGGCTGGAGAC	Pst I
108	Forward	AAAGAATTC-ATGTTGCCGGGCTTCAACCG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
108a	Forward	AAA AAAGAATTC-GGTAACACATTCCGGCAGCTTAGACGGTGG	Eco RI
	Reverse	AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Pst I
109	Forward	AAAGAATTC-ATGTATTATCGCCGGGTATGGG	Eco RI
	Reverse	AAACTGCAG-CTAGCCCAAAGATTTGAAGTGTC	Pst I
111	Forward	CGCGGATCCCATATG-TGTTCCGAACAAACCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGAGCAGTTTTTCAA	XhoI
114	Forward	CGCGGATCCCATATG-GCTTCCATCACTTCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATCCGCGAAATCGTC	XhoI
117	Forward	AAAAAAGGTACC-ATGGTCGAAGAACTGGAAGTCTG	Kpn I
	Reverse	AAACTGCAG-TTAAAGCCGGGTAACGCTCAATAC	Pst I
118	Forward	AAAGTCGACATGTGTGAGTTCAAGGATATTATAAG	Sal I
	Reverse	AAAGCATGC-CTATTTTTTGTGTAATAATCAAATC	Sph I
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC	XhoI
122	Forward	CGCGGATCCCATATG-GTCATGATTAAATCCGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATCTTGGTAGATTGGATTT	XhoI
125	Forward	AAAGAATTC-ATGTCGGGCAATGCCTCCTCTCC	Eco RI
	Reverse	AAACTGCAG-TCACGCCGTTTCAAGACG	Pst I
125a	Forward	AAAAAAGAATTC-ACGGCAGGCAGCACCGCCGCACAGGTTTC	Eco RI
	Reverse	AAAAAACTGCAG- TTATTTTGCCACGTCGGTTTCTCCGGTGAACAACGC	Pst I
126	Forward	CGCGGATCCCATATG-CCGTCTGAAACCC	BamHI-

	Reverse	CCCGCTCGAG-ATATTCCGCCGAATGCC	NdeI
			XhoI
127	Forward	AAAGAATTC-ATGGAAATATGGAATATGTTGGACACTTG	Eco RI
	Reverse	AAACTGCAG-TTAAAGTGTTCGGAGCCGGC	Pst I
127a	Forward	AAAAAAGAATTC-AAGGAACTGATTATGTGTCTGTCTCGGG	Eco RI
	Reverse	AAACTGCAG-TTAAAGTGTTCGGAGCCGGC	Pst I
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GACCGCGTTGTGCGAAA	XhoI
130	Forward	CGCGGATCCCATATG-AAACAACCTCCGCGA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GAATTTTGCACCGGATTG	XhoI
132	Forward	AAAGAATTC-ATGGAACCTTCAAACCTTAATTTG	Eco RI
	Reverse	AAAAAACTGCAG-TCACCATGTCTGGCATTGAAAAAC	Pst I
134	Forward	CGCGGATCCCATATG-TCCCAAGAAATCCTC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CAGTTTGACCGAATGTTC	XhoI
135	Forward	CGCGGATCCCATATG-AAATACAAAAGAATCGTATT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAATTCGGTCAGAAGCAGG	XhoI
137	Forward	AAAAAAGGTACC-ATGATTACCCATCCCAATTCGATCC	Kpn I
	Reverse	AAAAAACTGCAG-TCAGTGCTGTTTTTTCATGCCGAA	Pst I
137a	Forward	AAAAAAGAATTC-GGCCGCAAACACGGCATCGGCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TTAAGCGGGATGACGCGGCAGCATACC	Pst I
138	Forward	AAAAAAGAAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
141	Forward	AAAGAATTC-ATGAGCTTCAAACCGATGCCGAAATCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAACAAGCCGTGAATCACGCC	Pst I
142	Forward	CGCGGATCCCATATG-CGTGCCGATTTTCATG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAACTGCTGCACATGGG	XhoI
143	Forward	AAAAAAGAATTC- ATGCTCAGTTTCGGCTTTCTCGGCGTTCAGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCCCGCCGTGTGTTTCTTTAAT	Pst I
144	Forward	AAAAAAGAATTC-GGTCTGATCGACGGGCGTGCCGTAAC	Eco RI
	Reverse	AAAAAATCTAGA-TCGGCATCGGCCGGCATATGTCCG	Xba I
146	Forward	AAAAAAGAATTC- CGCCAAGTCGTCATTGACCACGACAAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGCATCGGCAAATAGGAACTGGG	Pst I
147	Forward	AAAAAAGAATTC-ACTGAGCAATCGGTGGATTTGAAAC	Eco RI
	Reverse	AAAAAATCTAGA-TTAGGTAAAGCTGCGGCCCATTTGCGG	Xba I
148	Forward	AAAAAAGAATTC- ATGGCGTTAAAAACATCAAACCTTGAACACGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGCCCTTCATACAGCCTTCGTTTTG	Xba I
149	Forward	CGCGGATCCCATATG-CTGCTTGACAACAAAGT	BamHI-



	Reverse	CCCGCTCGAG-AACTTCACGTTACGCC	NdeI
150	Forward	CGCGGATCCCATATG-CAGAACACAAATCCG	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-ATAAACATCACGCTGATAGC	NdeI
151	Forward	AAAAAAGAATTC- ATGAAACAAATCCGCAACATCGCCATCATCGC	XhoI
			Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAGCTTTTAAAGTGGCGGCG	Pst I
152	Forward	AAAAAAGAATTC- ATGAAAAACAAAACCAAAGTCTGGGACCTCCC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGGACAGGAGCAGGATGGCGGC	Pst I
153	Forward	AAAAAAGAATTC-ATGGCGTTTGCTTACGGTATGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTCATGTTTTCCGTTTCATT	Pst I
153a	Forward	AAAAAAGAATTC-CGGACTTCGGTATCGGTTCCCCAGCATTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCCGACGAAATACTCAGACTTTTCGG	Pst I
154	Forward	CGCGGATCCCATATG-ACTGACAACAGCCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCGGCTTCCTTTTCGGG	XhoI
155	Forward	AAAAAAGAATTC-ATGAAAATCGGTATCCCACGCGAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCCTTTCTTAAACATATTCAGCAT	Pst I
156	Forward	AAAAAAGAATTC-GCACAGCAAAACGGTTTTGAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCAGCCGCGACAAACAGCCC	Pst I
157	Forward	CGCGGATCCCATATG-AGGAACGAGGAAAAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAACACAATATCCCCGC	XhoI
158	Forward	AAAAAAGAATTC-GCGGAGCAGTTGGCGATGGCAAATTCTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTATCCACAGAGATTGTTTCCCAGTTC	Xba I
160	Forward	CGCGGATCCCATATG-GACATTCTGGACAAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTTTGCCCGCCTTCTTT	XhoI
163	Forward	AAAAAAGGTACC-ACCGTGCCGGATCAGGTGCAGATGTG	Kpn I
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATCCACCTGCTCGTG	Xba I
163a	Forward	AAAAAAGAATTC-CGGCTGGTGCAGATAATGAGCCAGAC	Eco RI
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATCCACCTGCTCGTG	Xba I
164	Forward	CGCGGATCCCATATG-AACCGGACTTATGCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGTTTCCGTCAAACCTGC	XhoI
165	Forward	CGCGGATCCGCTAGC-GCTGAAGCGACAGACG	BamHI-
			NheI
	Reverse	CCCGCTCGAG-AATATCCAATACTTTTCGCG	XhoI
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC	XhoI
209	Forward	CGCGGATCCCATATG-CTGCGGCATTTAGGA	BamHI-
			NdeI

	Reverse	CCCGCTCGAG-TACCCCTGAAGGCAAC	XhoI
211	Forward	AAAAAAGAATTC-ATGTTGCGGGTTGCTGCTGC	Eco RI
	Reverse	AAAAAACTGCAG-CTATCCTGCGGATTGGCATTGAAA	Pst I
212	Forward	CGCGGATCCCATATG-GACAATCTCGTATGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGGGTTAGATCCTTCC	XhoI
215	Forward	CGCGGATCCCATATG-GCATGGTTGGGTCTGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATATCTTTTGTATCATAAATC	XhoI
216	Forward	CGCGGATCCCATATG-GCAATGGCAGAAAACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TACAATCCGTGCCGCC	XhoI
217	Forward	CGCGGATCCCATATG-GCGGATGACGGTGTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACCCCGAATATCGAATCC	XhoI
218	Forward	CGCGGATCCCATATG-GTCGCGGTCTGATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TAACTCATAGAATCCTGC	XhoI
219	Forward	CGCGGATCCGCTAGC-ACGGCAAGGTTAAG	BamHI- NheI
	Reverse	CCCGCTCGAG-TTTAAACCATCTCCTCAAAAC	XhoI
223	Forward	CGCGGATCCCATATG-GAATTCAGGCACCAAGTA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCTTCCCGCGTGTG	XhoI
225	Forward	CGCGGATCCCATATG-GACGAGTTGACCAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTTCAGAAAGCGGGAC	XhoI
226	Forward	AAAGAATTC-CTTGCGATTATCGTGCGCACGCG	Eco RI
	Reverse	AAACTGCAG-TCAAAATCCCAAAACGGGGAT	Pst I
228	Forward	CGCGGATCCCATATG-TCGCAAGAAGCCAAACAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI
229	Forward	CGCGGATCCCATATG-CAAGAGGTTTTGCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACACAATATAGCGGATGAAC	XhoI
230	Forward	CGCGGATCCCATATG-CATCCGGGTGCCGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAGTTTGGCGGCTTCGG	XhoI
232	Forward	AAAAAAGAATTC-ATGTACGCTAAAAAAGGCGGTTTGGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGGTTTTTCTGATTGCCGCCG	Pst I
232a	Forward	AAAAAAGAATTC-GCCAAGGCTGCCGATACACAAATTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACATTGTCGTTGCCGCCAGATG	Pst I
233	Forward	CGCGGATCCCATATG-GCGGACAAACCCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGGCATTGAGCAG	XhoI
234	Forward	CGCGGATCCCATATG-GCCGTTTCACTGACCG	BamHI-

	Reverse	GCCCAAGCTT-ACGGTTGGATTGCCATG	NdeI
235	Forward	CGCGGATCCCATATG-GCCTGCCAAGTTCAAA	Hind III
	Reverse	CCCGCTCGAG-TTTGGGCTGCTCTTC	BamHI-
236	Forward	CGCGGATCCCATATG-GCGCGTTTCGCCTT	NdeI
	Reverse	CCCGCTCGAG-ATGGGTCGCGCGCCGT	XhoI
238	Forward	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCCG	BamHI-
	Reverse	CCCGCTCGAG-TTTGTCTAAGTTCCTGATATG	NdeI
239	Forward	CCGGAATTCTACATATG-CTCCACCATAAAGGTATTG	XhoI
	Reverse	CCCGCTCGAG-TGGTGAAGAGCGGTTTAG	EcoRI-
240	Forward	CGCGGATCCCATATG-GACGTTGGACGATTTC	NdeI
	Reverse	CCCGCTCGAG-AAACGCCATTACCCGATG	XhoI
241	Forward	CCGGAATTCTACATATG-CCAACACGTCCAAC	BamHI-
	Reverse	CCCGCTCGAG-GAATGCGCCTGTAATTAATC	NdeI
242	Forward	CGCGGATCCCATATG-ATCGGCAAACCTGTTG	XhoI
	Reverse	GCCCAAGCTT-ACCGATACGGTCGCAG	BamHI-
243	Forward	CGCGGATCCCATATG-ACGATTTTTTCGATGCTGC	NdeI
	Reverse	CCCGCTCGAG-CGACTTGGTTACCGCG	XhoI
244	Forward	CGCGGATCCCATATG-CCGTCTGAAGCCC	BamHI-
	Reverse	CCCGCTCGAG-TTTTTTCGGTAGGGGATTT	NdeI
246	Forward	CGCGGATCCCATATG-GACATCGGCAGTGC	XhoI
	Reverse	CCCGCTCGAG-CCCGCGCTGCTGGAG	BamHI-
247	Forward	CGCGGATCCCATATG-GTCGGATCGAGTTAC	NdeI
	Reverse	CCCGCTCGAG-AAGTGTTCTGTTTGCGCA	XhoI
248	Forward	CGCGGATCCCATATG-CGAAACAGAACACT	BamHI-
	Reverse	CCCGCTCGAG-CTCATCATTATTGCTAACA	NdeI
249	Forward	CGCGGATCCCATATG-AAGAATAATGATTGCTTC	XhoI
	Reverse	CCCGCTCGAG-TTCCCGACCTCCGAC	BamHI-
251	Forward	CGCGGATCCCATATG-CGTGCTGCGGTAGT	NdeI
	Reverse	CCCGCTCGAG-TACGAAAGCCGGTCGTG	XhoI
253	Forward	AAAAAAGAATTC-ATGATTGACAGGAACCGTATGCTGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I

253a	Forward	AAAAAAGAATTC-AAAATCCTTTTGAAAACAAGCGAAAACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I
254	Forward	AAAAAAGAATTC-ATGTATACAGGCGAACGCTTCAATAC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATTACGTAACCGTACACGCTGAC	Xba I
255	Forward	CGCGGATCCCATATG-GCCGCGTTGCGTTAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ATCCGCAATACCGACCAG	XhoI
256	Forward	CGCGGATCCGCTAGC-TTTTAACACCGCCGGAC	BamHI-
			NheI
	Reverse	CCCGCTCGAG-ACGCCTGTTTGTGCGG	XhoI
257	Forward	CGCGGATCCCATATG-GCGGTTTCTTTCCTG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GCGCGTGAATATCGCG	XhoI
258	Forward	AAAAAAGAATTC-GATTATTTCTGGTGGATTGTTGCGTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
258a	Forward	AAAAAAGAATTC-GCGAAGGCGGTGGCCCAAGGCGA	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
259	Forward	CGCGGATCCCATATG-GAAGAGCTGCCTCCG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GGCTTTTCCGGCGTTT	XhoI
260	Forward	CGCGGATCCCATATG-GGTGCGGGTATGGT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AACAGGGCGACACCCT	XhoI
261	Forward	AAAAAAGAATTC-CAAGATACAGCTCGGGCATTTCG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCAACAAGCCTTGGTCACT	Pst I
263	Forward	CGCGGATCCCATATG-GCACGTTTAACCGTA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GGCGTAAGCCTGCAATT	XhoI
264	Forward	AAAAAAGGTACC-GCCGACGCAGTGGTCAAGGCAGAA	Kpn I
	Reverse	AAACTGCAG-TCAGCCGGCGGTCAATACCGCCCG	Pst I
265	Forward	AAAAAAGAATTC-GCGGAGGTCAAGAGAAGGTGTTTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGAATACGTCGTCAAATGGG	Pst I
266	Forward	AAAGAATTC-CTCATCTTTGCCAACGCCCCCTTC	Eco RI
	Reverse	AAACTGCAG-CTATTCCTGTTGCGCGTGTGCCA	Pst I
267	Forward	AAAGAATTC-TTCTCCGATTTCGATGTTAATCG	Eco RI
	Reverse	AAACTGCAG-TTAGTAAAAACCTTTCTGCTTGGC	Pst I
269	Forward	AAAGAATTC-TGCAAACCTTGCGCCACGTGCCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
269a	Forward	AAAAAAGAATTC-GACTTTATCCAAAACACGGCTTCGCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
270	Forward	AAAGAATTC-GCCGTCAAGCTCGTTTTGTTGCAATG	Eco RI
	Reverse	AAACTGCAG-TTATTCGGCGGTAAATGCCGTCTG	Pst I
271	Forward	CGCGGATCCCATATG-CCTGTGTGCAGCTCGAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCCCAGCCCCGTGGAG	XhoI

272	Forward	AAAGAATTC-ATGACCCCAAAGGAAGAACTGTTGCG	Eco RI
	Reverse	AAACTGCAG-TCAGAGCAGTTCCAAATCGGGGCT	Pst I
273	Forward	AAAGAATTC-ATGAGTCTTCAGGCGGTATTTATATACCC	Eco RI
	Reverse	AAACTGCAG-TTACGCGTAAGAAAAAACTGC	Pst I
274	Forward	CGCGGATCCCATATG-ACAGATTGGTTACGGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGCTTTCAGTATTATTGAA	XhoI
276	Forward	AAAAAAGAATTC- ATGATTTTGCCGTCGTCCATCACGATGATGCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACACCACCATCGGCGAATTTATGGC	Pst I
277	Forward	AAAAAAGAATTC-ATGCCCCGCTTTGAGGACAAGCTCGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
277a	Forward	AAAAAAGAATTC-GGGGCGGCGGCTGGGTGGACGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
278	Forward	AAAAAAGGTACC-GTCAAAGTTGTATTAATCGGGCCTTTGCC	Kpn I
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
278a	Forward	AAAAAAGAATTC-AAAACTCTCCTAATTCGTCATAGTCG	Eco RI
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
279	Forward	CGCGGATCCCATATG-TTGCTGCAATCACGATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA	XhoI
280	Forward	AAAAAAGGTACC-GCCCCCTGCCGTTGTAAACCAG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTGCTTCATCGCGTTGGTCAAGGC	Pst I
281	Forward	AAAAAAGAATTC-GCACCCGTCGGCGTATTCCTCGTCATGCG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
281a	Forward	AAAAAAGAATTC-TCCTACCACATCGAAATTCCTTCCGG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
282	Forward	AAAAAAGAATTC-CTTTACCTTGACCTGACCAACGGGCACAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACCTGCCAGTTGCGGGAATATCGT	Pst I
283	Forward	CGCGGATCCCATATG-GCCGTCTTTACTTGGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGGCAGTATTTGTTTACG	XhoI
284	Forward	CGCGGATCCCATATG-TTGCCTGCAAAAGAATCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCGACTTTGCAAAAACCTG	XhoI
286	Forward	CGCGGATCCCATATG-GCCGACCTTCCGAAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAAGCGCGTTCCCAAG	XhoI
287	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG	EcoRI- NheI
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC	XhoI
288	Forward	CGCGGATCCCATATG-CACACCGGACAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGTATCAAAGACTTGCGT	XhoI
290	Forward	CGCGGATCCCATATG-GCGGTTTGGGGCGGA	BamHI-

	Reverse	CCCGCTCGAG-TCGGCGCGGCGGGC	NdeI
292	Forward	CGCGGATCCCATATG-TGCGGGCAAACGCCC	XhoI
			BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTGATTTTTGCGGATGATTT	XhoI
294	Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	Pst I
295	Forward	CGCGGATCCCATATG-AACCGGCCGGCCTCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CGATATTTGATTCCGTTGC	XhoI
297	Forward	AAAAAAGAATTC-GCATAATTGCTTCGACAGAGAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCGATTGCGACACGGT	Pst I
298	Forward	AAAAAAGAATTC-CTGATTGCCGTGTGGTTCAGCCAAAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGGCTGTGTAC GTGATGGTTGCGT	Pst I
299	Forward	CGCGGATCCGCTAGC-CTACCTGTCGCCTCCG	BamHI-
			NheI
	Reverse	CCCGCTCGAG-TTGCCTGATTGCAGCGG	XhoI
302	Forward	AAAAAAGAATTC-ATGAGTCAAACCGATACGCAACG	Eco RI
	Reverse	AAAAAACTGCAG-TTAAGGTGCGGGATAGAATGTGGGCGC	Pst I
305	Forward	AAAAAAGGTACC-GAATTTTTACCGATTTCCAGCACCGGA	Kpn I
	Reverse	AAAAAACTGCAG-TCATTCCCAACTTATCCAGCCTGACAG	Pst I
305a	Forward	AAAAAAGGTACC-TCCCGTTCGGGCAGTACGATTATGGG	Kpn I
	Reverse	AAAAAACTGCAG-TTACAAACCGACATCATGCAGGGTGAA	Pst I
306	Forward	CGCGGATCCCATATG-TTTATGAACAAATTTTCCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CCGCATCGGCAGAC	XhoI
308	Forward	CGCGGATCCCATATG-TTAAATCGGGTATTTTATC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ATCCGCCATTCCCTGC	XhoI
311	Forward	AAAAAAGGTACC-ATGTTTCAGTTTTGGCTGGGTGTTT	Kpn I
	Reverse	AAACTGCAG-ATGTTTCATATTCCTGCCTTCGGC	Pst I
312	Forward	AAAAAAGGTACC-ATGAGTATCCCATCCGGCGAAATT	Kpn I
	Reverse	AAACTGCAG-TCAGTTTTTTCATCGATTGAACCGG	Pst I
313	Forward	AAAAAAGAATTC-ATGGACGACCCGCGCACCTACGGATC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGCGGCTGCCGCCGATTTTGCT	Pst I
401	Forward	CGCGGATCCCATATG-AAGGCGGCAACACAGC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CCTTACGTTTTTCAAAGCC	XhoI
402	Forward	AAAAAAGAATTC-GTGCCTCAGGCATTTTCATTTACCCTTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Xba I
402a	Forward	AAAAAAGAATTC-AGGCTGATTGAAAACAAACACGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Xba I
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCTTGTCTATG	XhoI

501	Forward	CGCGGATCCCATATG-GCAGGCGGAGATGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGTGTGATGTTACCC	XhoI
502	Forward	CGCGGATCCCATATG-GTAGACGCGCTTAAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCTGCATGGCGGCG	XhoI
503	Forward	CGCGGATCCCATATG-TGTTCGGGGAAAGGCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCGCGCATTCCTCGCA	XhoI
504	Forward	CGCGGATCCCATATG-AGCGATATTGAAGTGACG	BamHI- NdeI
	Reverse	GCCCAAGCTT-TGATTCAAGTCCTTGCCG	HindIII
505	Forward	CGCGGATCCCATATG-TTCGTTTACAATTGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGTTTTATAGCGG	XhoI
510	Forward	CGCGGATCCCATATG-CCTTCGCGGACAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCACTGGCAGCG	XhoI
512	Forward	CGCGGATCCCATATG-GGACATGAAGTAACGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGAATAGCCTTTGACG	XhoI
515	Forward	CGCGGATCCCATATG-GAGGAAATAGCCTTCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATGCCGCAAAGCATC	XhoI
516	Forward	CGCGGATCCCATATG-TGTACGTTGATGTTGTGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGCGGGCGGCATC	XhoI
517	Forward	CGCGGATCCCATATG-GGTAAAGGTGTGGAAATA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGCGCCCAGCCGT	XhoI
518	Forward	AAAGAATTC-GCTTTTTTACTGCTCCGACCGGAAGG	Eco RI
	Reverse	AAACTGCAG-TCAAATTTGAGACTCTGCCAC	Pst I
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC	XhoI
520	Forward	CGCGGATCCCATATG-CCTGCGCTTCTTTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATATTTACATTTGAGTCGGC	XhoI
521	Forward	CGCGGATCCCATATG-GCCAAAATCTATACCTGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATACGCCCCAGTTCC	XhoI
522	Forward	CGCGGATCCCATATG-ACTGAGCCGAAACAC	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTCTGATTTCAAATCGGCA	HindIII
523	Forward	CGCGGATCCCATATG-GCTCTGCTTTCCGCG	BamHI- NdeI

	Reverse	CCCGCTCGAG-AGGGTGTGTGATAATAAGAAG	XhoI
525	Forward	CGCGGATCCCATATG-GCCGAAATGGTTCAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCCGTGCATATCATAAA	XhoI
527	Forward	AAAGAATTC-TTCCCTCAATGTTGCCGTTTTTCG	Eco RI
	Reverse	AAACTGCAG-TTATGCTAAACTCGAAACAAATTC	Pst I
529	Forward	CGCGGATCCGCTAGC-TGCTCCGGCAGCAAAAC	BamHI- NheI
	Reverse	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
530	Forward	CGCGGATCCCATATG-AGTGCGAGCGCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGACCGACTGATTCCG	XhoI
531	Forward	AAAAAAGAATTC-TATGCCGCCGCTACCAAATCTACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACAGCGCCGTGCCGACGACAAG	Pst I
532	Forward	AAAAAAGAATTC-ATGAGCGGTCAGTTGGGCAAAGGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCGAAGTCCTCGGTATCAAA	Pst I
532a	Forward	AAAAAAGAATTC-TTGGGTGTCGCGTTTGAGCCGGAAGT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCGAAGTGGTCGGTATCAAA	Pst I
535	Forward	AAAGAATTC-ATGCCCTTTCGTTTTCAGAC	Eco RI
	Reverse	AAACTGCAG-TCAGACGACCCCGCCTTCCCC	Pst I
537	Forward	CGCGGATCCCATATG-CATACCCAAAACCAATCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCCTGCAAATAAAGGGTT	XhoI
538	Forward	CGCGGATCCCATATG-GTCGAGCTGGTCAAAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGGCATTTCGGTTTCGTC	XhoI
539	Forward	CGCGGATCCGCTAGC-GAGGATTTGCAGGAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TACCAATGTCCGCAAATC	XhoI
542	Forward	AAAGAATTC-ATGCCGTCTGAAACCGTGTC	Eco RI
	Reverse	AAACTGCAG-TTACCGCGAACC GGTCAGGAT	Pst I
543	Forward	AAAAAAGAATTC-GCCTTCGATGGCGACGTTGTAGGTAC	Eco RI
	Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
543a	Forward	AAAAAAGAATTC-GGCAAACTCGTCATGAATTTGC	Eco RI
	Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
544	Forward	AAAGAATTC-GCGCCCGCCTTCTCCCTGCCCGACCTGCACGG	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGGTATCGAT	Pst I
544a	Forward	AAAAAAGAATTC- GCAATGACTATAAAAAACAAAACTTCCAAGTACTTGC	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGGTATCGAT	Pst I
547	Forward	AAAGAATTC-ATGTTCTGTAGATAACGGATTTAATAAAAC	Eco RI
	Reverse	AAACTGCAG-TTAACAACAAAAACAAACCGCTT	Pst I
548	Forward	AAAGAATTC-GCCTGCAAACCTCAAGACAACAGTGCGGC	Eco RI



	Reverse	AAACTGCAG-TCAGAGCAGGGTCCTTACATCGGC	Pst I
550	Forward	AAAAAAGTCGAC- ATGATAACGGACAGGTTTCATCTCTTTTCATTTTCC	Sal I
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAATCCCC	Pst I
550a	Forward	AAAAAAGAATTC-GTAAATCACGCCTTTGGAGTCGCAAACGG	Eco RI
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAATCCCC	Pst I
552	Forward	AAAAAAGAATTC-TTGGCGCGTTGGCTGGATAC	Eco RI
	Reverse	AAACTGCAG-TTATTTCTGATGCCTTTTCCCAAC	Pst I
554	Forward	CGCGGATCCCATATG-TCGCCCGCGCCCAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGCCCTGTCAGACAC	XhoI
556	Forward	AAAGAATTC-GCGGGCGGTTTTGTTGGACATCCCG	Eco RI
	Reverse	AAACTGCAG-TTAACGGTGCGGACGTTTCTGACC	Pst I
557	Forward	CGCGGATCCCATATG-TGCGGTTTCCACCTGAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTCCGCCTTCAGAAAGG	XhoI
558	Forward	AAAGAATTC-GAGCTTTATATGTTTCAACAGGGGACGGC	Eco RI
	Reverse	AAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
558a	Forward	AAAAAAGAATTC-ATTAGATTCTATCGCCATAAACAGACGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
560	Forward	AAAAAAGAATTC- TCGCCTTTCCGGGACGGGGCGCACAAGATGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGCGGTTTTCAGACGGCATTTTGGC	Pst I
561	Forward	CCGGAATTCTACATATG-ACTGCGCAGCCCGT	EcoRI- NdeI
	Reverse	CCCGCTCGAG-TTTCAAGCTTTCTTCAGATG	XhoI
562	Forward	CGCGGATCCCATATG-GCAAGCCCGTCGAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGACCAACTCCAACCTCGT	XhoI
565	Forward	CGCGGATCCCATATG-AAGTCGAGCGCGAAATAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCATTGATCGGCGGC	XhoI
566	Forward	CGCGGATCCCATATG-GTCGGTGGCGAAGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCATGGGCGAAGTCA	XhoI
567	Forward	CCGGAATTCTACATATG-AGTGCGAACATCCTTG	EcoRI- NdeI
	Reverse	CCCGCTCGAG-TTTCCCCGACACCCTCG	XhoI
568	Forward	CGCGGATCCCATATG-CTCAGGGTCAGACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGCGGCGTTTCAG	XhoI
569	Forward	AAAAAAGAATTC-CTGATTGCCTTGTTGGGAATATGCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATGCATAGACGCTGATAACGGCAAT	Pst I
570	Forward	CGCGGATCCCATATG-GACACCTTCCAAAAAATCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGGCGTTTCATTTCTTT	XhoI

571	Forward	AAAAAAGAATTC- ATGGGTATTGCCGGCGCCGTAAATGTTTTGAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATGGCCGACGCGCGGCTACCTGACG	Pst I
572	Forward	CGCGGATCCCATATG-GCGCAAAAAGGCAAAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCAGTGTGCCGATA	XhoI
573	Forward	CGCGGATCCCATATG-CCCTGTTTGTGCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGGTGTCAATTCGCC	XhoI
574	Forward	CGCGGATCCCATATG-TGGTTTGCCGCCCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACTTCGATTTTATTCGGG	XhoI
575	Forward	CGCGGATCCCATATG-GTTTCGGGCGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCCGAATCTGAACAG	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCGCATCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC	XhoI
577	Forward	CGCGGATCCCATATG-GAAAGGAACGGTGTATTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGCTGTTTGGTAGATTCTG	XhoI
578	Forward	CGCGGATCCCATATG-AGAAGGTTCTGACAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCAACGCCTCCACG	XhoI
579	Forward	CGCGGATCCCATATG-AGATTGGGCGTTTCCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGAATTGATGATGTGTATGT	XhoI
580	Forward	CGCGGATCCCATATG-AGGCAGACTTCGCCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACTTCCCCCGAAGTG	XhoI
581	Forward	CGCGGATCCCATATG-CACTTCGCCCAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCCGTTTGGCTTTGG	XhoI
582	Forward	AAAAAAGAATTC-TTTGGAGAGACCGCGCTGCAATGCGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATGCCGTCCCAGTCGTTGAA	Xba I
583	Forward	AAAAAAGAATTC-ACTGCCGGCAATCGACTGCATAATCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGAGGTCAATATGATGAAATTG	Pst I
584	Forward	AAAAAAGAATTC- GCGGCTGAAGCATTGAATTACAATATTGTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGAACTGAACCGTCCCATTGACGCT	Pst I
585	Forward	AAAAAAGGTACC-TCTTTCTGGCTGGTGCAGAACACCCTTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTTCGCACTTTTTTCTGTTTGGGA	Pst I
586	Forward	CGCGGATCCCATATG-GCAGCCCATCTCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTCAGCGAATCAAGTTTC	XhoI
587	Forward	CGCGGATCCCATATG-GACCTGCCCTTGACGA	BamHI-

	Reverse	CCCGCTCGAG-AAATGTATGCTGTACGCC	NdeI
			XhoI
588	Forward	AAAAAAGAATTC-GCCGTCCTGACTTCCTATCAAGAACCAGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTGTTTTTGGGCAGTTTCACTTC	Pst I
589	Forward	AAAAAAGAATTC-	Eco RI
		ATGCAACAAAAAATCCGTTTCCAAATCGAAGG	
	Reverse	AAAAAACTGCAG-CTAATCGATTTTACCCGTTTCAGGCG	Pst I
590	Forward	AAAAAAGAATTC-ATGAAAAAACCTTTGATTTTCAGTTGCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCTGCGGCTCTGAAACCAT	Pst I
591	Forward	AAAAAAGAATTC-CACTACATCGTTGCCAGATTGTGCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAACCAGCAGCCGGGTAACGTCGTT	Pst I
592a	Forward	AAAAAAGAATTC-CGCGATTACACCGCCAAGCTGAAAATGGG	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAAACGTCGGATTTGATACG	Pst I
593	Forward	CGCGGATCCGCTAGC-CTTGAACCTGAACGGACTC	BamHI-
			NheI
	Reverse	CCCGCTCGAG-GCGGAAGCGGACGATT	XhoI
594a	Forward	AAAAAAGAATTC-GGTAAGTTCGCCGTTTCAGGCCTTTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCGCCGTTTCTGACACTCGCG	Pst I
595	Forward	AAAAAAGAATTC-TGCCAGCCGCCGGAGGCGGAGAAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCAAGCCGAGTATGCCGCG	Pst I
596	Forward	CGCGGATCCCATATG-TCCCAACAATACGTC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ACGCGTTACCGGTTTGT	XhoI
597	Forward	CGCGGATCCCATATG-CTGCTTCATGTCAGC	BamHI-
			NdeI
	Reverse	GCCCAAGCTT-ACGTATCCAGCTCGAAG	HindIII
601	Forward	CGCGGATCCCATATG-ATATGTTCCCAACCGGCAAT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAAACAATCCTCAGGCAC	XhoI
602	Forward	CGCGGATCCGCTAGC-TTGCTCCATCAATGC	BamHI-
			NheI
	Reverse	CCCGCTCGAG-ATGCAGCTGCTAAAAGCG	XhoI
603	Forward	AAAAAAGAATTC-CTGTCCTCGCGTAGGCGGGGACGGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAGATGCCGGCAAGTTCGGC	Pst I
604	Forward	CGCGGATCCGCTAGC-CCCGAAGCGCACTT	BamHI-
			NheI
	Reverse	CCCGCTCGAG-GACGGCATCTGCACGG	XhoI
606a	Forward	AAAAAAGAATTC-CGCGAATACCGCGCCGATGCGGGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAGCGATTTGAGGCGGGCGATACG	Pst I
607	Forward	AAAAAAGAATTC-ATGCTGCTCGACCTCAACCGCTTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCCTTATGCGATCTGAC	Pst I
608	Forward	AAAAAAGAATTC-ATGTCCGCCCTCCTCCCATCATCAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTCTATCCAAATGTCGCGTTC	Pst I
609	Forward	CGCGGATCCCATATG-GTTGTGGATAGACTCG	BamHI-
			NdeI

	Reverse	CCCGCTCGAG-CTGGATTATGATGTCTGTC	XhoI
610	Forward	CGCGGATCCCATATG-ATTGGAGGGCTTATGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGCTTCAACATCTTTGCC	XhoI
611	Forward	CGCGGATCCCATATG-CCGTCTCAAACGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACGACTTTGAACGCGCAA	XhoI
613	Forward	CGCGGATCCCATATG-TCGCGTTCGAGCCG3	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCCTGTAAAATAAGCGGC	XhoI
614	Forward	CGCGGATCCCATATG-TCCGTCTGAGCGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCATACTGCGGCGTTC	XhoI
616	Forward	AAAAAAGAATTC-ATGTCAAACACAATCAAAATGGTTGTCGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAGTCCGGGCGGCAGGCAGCTCG	Xba I
619a	Forward	AAAAAAGAATTC-GGGCTTCTCGCCGCTTGGCTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TCATTTTTTGTGTTTTAAAACGAGATA	Pst I
622	Forward	CGCGGATCCCATATG-GCCGCCCTGCCTAAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGTCCAAATGATAAATCTG	XhoI
624	Forward	CGCGGATCCCATATG-TCCCCGCGCTTTTACCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGATTCGGGCCTGCGC	XhoI
625	Forward	CGCGGATCCCATATG-TTTGCAACCAGGAAAATG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCAAATTACCGCCTT	XhoI
627a	Forward	AAAAAAGAATTC-AAAGCAGGCGAGGCAGGCGCGCTGGG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGAATGAAACAGGGTACCCGTCATCAAGGC	Pst I
628	Forward	AAAAAAGGTACC-GCCTTACAAACATGGATTTTGC GTTC	Kpn I
	Reverse	AAAAAACTGCAG-CTACGCACCTGAAGCGCTGGCAA	Pst I
629a	Forward	AAAAAAGAATTC-GCCACCTTTATCGCGTATGAAAACGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACAACACCGCCGTCCGGTTCAAACC	Pst I
630a	Forward	AAAAAAGAATTC-GCGGCTTTGGGTATTTCTTTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGAGACTTCGCCAATGGAGCCGGG	Pst I
635	Forward	AAAAAAGAATTC- ATGACCCAGCGACGGGTCTGGCAAGCAAAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATCCACTATAATCCTGTTGCT	Pst I
638	Forward	AAAAAAGAATTC-ATGATTGGCGAAAAGTTTATCGTAGTTGG	Eco RI
	Reverse	AAAAAACTGCAG-TCACGAACCGATTATGCTGATCGG	Pst I
639	Forward	CGCGGATCCCATATG-ATGCTTTATTTTGTTCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCGCGGCTGCCGAC	XhoI
642	Forward	CGCGGATCCCATATG-CGGTATCCGCCGCAAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGATTGCGGGGCATTA	XhoI

643	Forward	CGCGGATCCCATATG-GCTTCGCCGTCGGCAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACCGAAAAACAGACCGC	XhoI
644	Forward	AAAAAAGAATTC- ATGCCGTCTGAAAGGTCGGCGGATTGTTGCCC	Eco RI
	Reverse	AAAAAATCTAGA-CTACCCGCAATATCGGCAGTCCAATAT	Pst I
645	Forward	AAAAAAGAATTC-GTGAACAGAGCAACACGTAAATCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGAGGAAACCGAAGACCAGGCCGC	Pst I
647	Forward	AAAAAAGAATTC-ATGCAAAGGCTCGCCGCAGACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGATTATCAGGGATATCCGGTAGAA	Pst I
648	Forward	AAAAAAGAATTC- ATGAACAGGCGCGACGCGCGGATCGAACG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCTGTGTGCTGATTGAATGCGAC	Pst I
649	Forward	AAAAAAGAATTC-GGTACGTCAGAACCCGCCACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGCGGAACTGCCGCCGTC	Pst I
650	Forward	AAAAAAGAATTC-ATGTCCAAACTCAAACCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCATGGCGGTCTGTTTT	Pst I
652	Forward	AAAAAAGGTACC- GCTGCCGAAGACTCAGGCCTGCCGCTTTACCG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTTGCCAGTTGGTAGAATGCGGC	Pst I
653	Forward	AAAAAAGAATTC-GCGGCTTTGCCGTAATTTTCATCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTATGCCGGTCTGGTTGCCGGCGGCGA	Pst I
656a	Forward	AAAAAAGAATTC-CGGCCGACGTCGTTGCGTCCTAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-CTACGATTTGCGGATTTCCACATCGT	Pst I
657	Forward	AAAAAAGAATTC-GCAGAATTTGCCGACCGCCATTTGTGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATAGGGACTGATGCAGTTTTTTTGC	Pst I
658	Forward	CGCGGATCCCATATG-GTGTCCGGAATTGTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCAGAATGTTTACCGTT	XhoI
661	Forward	AAAAAAGAATTC- ATGCACATCGGCGGCTATTTTATCGACAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCACGACGTGTCTGTTCCCGTCGGGC	Pst I
663	Forward	CGCGGATCCCATATG-TGTATCGAGATGAAATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTAAAAATCGGGGCTGC	XhoI
664	Forward	CGCGGATCCCATATG-GCGGCTGGCGCGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATCGAGTTTTACACCAC	XhoI
665	Forward	AAAAAAGAATTC-ATGAAATGGGACGAAACGCGCTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAAAATTTGCCGACGATTC	Pst I
666	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
667	Forward	AAAAAAGAATTC- CCGCATCCGTTTGATTTCCATTTTCGTATTCGTCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATGACACAATAGGCGCAAGTC	Pst I

669	Forward	AAAAAAGAATTC-ATGCGCCGCATCATTAAAAAACACCAGCC	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTATCCGTTTGATGTCGGC	Pst I
670a	Forward	AAAAAAGAATTC-AAAAACGCTTCGGGCGTTTCGTCTTC	Eco RI
	Reverse	AAAAAACTGCAG- TTAGGAGCTTTTGGAACGCGTCGGACTGGC	Pst I
671	Forward	CGCGGATCCCATATG-ACCAGCAGGGTAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCAACTATAAAAAACGCAAG	XhoI
672	Forward	CGCGGATCCCATATG-AGGAAAATCCGCACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGGGATAGGCGGTTG	XhoI
673	Forward	AAAAAAGAATTC-ATGGATATTGAAACCTTCCTTGCAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAACCCAGCTCGCGCAGGAA	Pst I
674	Forward	AAAAAAGAATTC-ATGAAAACAGCCCGCGCCGTTCCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACGGCGTTTGGGCTCGTCGGG	Pst I
675	Forward	CGCGGATCCCATATG-AACACCATCGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTCTTCGTCTTCAAACCTGT	XhoI
677a	Forward	AAAAAAGAATTC-AGACGGCATTCCCGATCAGTCGATTTTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGTATGCGCGAAATCGACCGCCGC	Pst I
680	Forward	CGCGGATCCGCTAGC-ACGAAGGGCAGTTCGG	BamHI- NheI
	Reverse	CCCGCTCGAG-CATCAAAAACCTGCCGC	XhoI
681	Forward	AAAAAAGAATTC-ATGACGACGCCGATGGCAATCAGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCGTCTTCCGCAAAAAACAGC	Pst I
683	Forward	CGCGGATCCCATATG-TGCAGCACACCGGACAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAGTTTTTTTCCGCATACG	XhoI
684	Forward	CGCGGATCCCATATG-TGCGGTACTGTGCAAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTCGACCATCTGTTGCG	XhoI
685	Forward	CGCGGATCCCATATG-TGTTTGCTTAATAATAAACATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTTTTTCCCCGCCGCA	XhoI
686	Forward	CGCGGATCCCATATG-TGCGGCGGTTTCGGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCCGATTCTGATGAAG	XhoI
687	Forward	CGCGGATCCCATATG-TGCGACAGCAAAGTCCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGCGCGGCTTTTTGTT	XhoI
690	Forward	CGCGGATCCCATATG-TGTTCTCCGAGCAAAGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TATTCGCCCCGTGTTTGG	XhoI
691	Forward	CGCGGATCCCATATG-GCCACGGCTTATATCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAGGCAGGAAGAAAG	XhoI

694	Forward	CGCGGATCCCATATG-TTGGTTTCCGCATCCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCTGCGTCCGGTGC	XhoI
695	Forward	CGCGGATCCCATATG-TTGCCTCAAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
696	Forward	CGCGGATCCCATATG-TTGGGTTGCCGGCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGATTGCCGCAATGATG	XhoI
700a	Forward	AAAAAAGAATTC-GCATCGACAGACGGTGTGTCGTGGAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCTACCGGCACGACTTCCAAACC	Pst I
701	Forward	CGCGGATCCCATATG-AAGACTTGTTTGATACTTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCCGACAACAGCCTC	XhoI
702	Forward	AAAAAAGAATTC-ATGCCGTGTTCCAAAGCCAGTTGGATTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACCCCATTCACCCGGAGAACCGA	Pst I
703	Forward	CGCGGATCCGCTAGC-CAAACGCTGGCAACCG	BamHI- NheI
	Reverse	CCCGCTCGAG-TTTTGCAGGTTTGATGTTTG	XhoI
704a	Forward	AAAAAAGAATTC-GCTTCTACCGGTACGCTGGCGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTTTGGCCGATAATATGGCGGGTGCG	Pst I
707	Forward	CGCGGATCCGCTAGC-GAAATTATTAACGATGCAGA	BamHI- NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGA	XhoI
708	Forward	CGCGGATCCGCTAGC-CCTTTTAAGCCATCCAAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TTGACCGGTGAGGACG	XhoI
710	Forward	CGCGGATCCCATATG-GAAACCCACGAAAAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACGGTTTCGGTCAG	XhoI
714	Forward	CGCGGATCCCATATG-AGCTATCAAGACATCTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGTAGGTAAATCGGAT	XhoI
716	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
718	Forward	CGCGGATCCCATATG-GAGCCGATAATGGCAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGCGGGCATGGTCTTGTC	XhoI
720	Forward	CGCGGATCCCATATG-AGCGGATGGCATACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTTGCATAGCTGTTGACCA	XhoI
723	Forward	CGCGGATCCCATATG-CGACCCAAGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATGCGAATCCGCCGCC	XhoI

725	Forward	CGCGGATCCCATATG-GTGCGCACGGTTAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGCTTATCCTTAAGGGTTA	XhoI
726	Forward	CGCGGATCCCATATG-ACCATCTATTTCAAAAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCGATGTTTAGCGTCC	XhoI
728	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	XhoI
729	Forward	CGCGGATCCCATATG-TGCACCATGATTCCCCA	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTTGTCGGTTTGGGTATC	HindIII
731	Forward	CGCGGATCCGCTAGC-GCCGTGCCGGAGG	BamHI- NheI
	Reverse	CCCGCTCGAG-ACGGGCGCGGCAG	XhoI
732	Forward	CCGGAATTCTACATATG-TCGAAACCTGTTTTTAAGAA	EcoRI- NdeI
	Reverse	CCCGCTCGAG-CTTCTTATCTTTTTTATCTTTC	XhoI
733	Forward	CGCGGATCCCATATG-GCCTGCGGCGGCAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGCTTGCCTCCTTTAC	XhoI
734	Forward	CGCGGATCCCATATG-GCCGATACTTACGGCTAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAGATTTTGAATCAAAGAG	XhoI
735	Forward	CGCGGATCCCATATG-AAGCAGCAGGCGGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTCCGTAGCCGAGGG	XhoI
737	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	XhoI
739	Forward	CGCGGATCCCATATG-GCAAAAAAACGAACA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAAGAGTTTGTGAGAATT	XhoI
740	Forward	CGCGGATCCCATATG-GCCAATCCGCCCGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAACGCGCCAAAATAGTG	XhoI
741	Forward	CGCGGATCCCATATG-TGCAGCAGCGGAGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI
743	Forward	CGCGGATCCCATATG-GACGGTGTTGTGCCTGTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTTACGGATCAAATTGACG	XhoI
745	Forward	CGCGGATCCCATATG-TTTTGGCAACTGACCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CAAATCAGATGCCTTTAGG	XhoI
746	Forward	CGCGGATCCCATATG-TCCGAAAACAAACAAAAC	BamHI-



	Reverse	CCCGCTCGAG-TTCATTCGTTACCTGACC	NdeI
747	Forward	CCGGAATTCTAGCTAGC-CTGACCCCTTGGG	XhoI
	Reverse	GCCCAAGCTT-TTTTGATTTTAATTGACTATAGAAC	EcoRI-
749	Forward	CGCGGATCCCATATG-TGCCAGCCGCCG	NheI
	Reverse	CCCGCTCGAG-TTCAAGCCGAGTATGC	HindIII
750	Forward	CGCGGATCCCATATG-TGTTGCCCCGAACCTG	BamHI-
	Reverse	CCCGCTCGAG-CTTTTTCCCCGCCGCAA	NdeI
758	Forward	CGCGGATCCCATATG-AACAATCTGACCGTGTT	XhoI
	Reverse	CCCGCTCGAG-TGGCTCAATCCTTTCTGC	BamHI-
759	Forward	CGCGGATCCGCTAGC-CGCTTCACACACACCAC	NdeI
	Reverse	CCCGCTCGAG-CCAGTTGTAGCCTATTTTG	XhoI
763	Forward	CGCGGATCCCATATG-CTGCCTGAAGCATGGCG	BamHI-
	Reverse	CCCGCTCGAG-TTCCGCAAATACCGTTTCC	NdeI
764	Forward	CGCGGATCCCATATG-TTTTCTCCGCCCTGA	XhoI
	Reverse	CCCGCTCGAG-TCGCTCCCTAAAGCTTTC	BamHI-
765	Forward	CGCGGATCCCATATG-TTAAGATGCCGTCCG	NdeI
	Reverse	CCCGCTCGAG-ACGCCGACGTTTTTTATTAA	XhoI
767	Forward	CGCGGATCCCATATG-CTGACGGAAGGGGAAG	BamHI-
	Reverse	CCCGCTCGAG-TTTCTGTACAGCAGGGG	NdeI
768	Forward	CGCGGATCCCATATG-GCCCCGCAAAAACCCG	XhoI
	Reverse	CCCGCTCGAG-TTTCATCCCTTTTTTGAGC	BamHI-
770	Forward	CGCGGATCCCATATG-TGCGGCAGCGGCGAA	NdeI
	Reverse	CCCGCTCGAG-GCGTTTGTGAGATTTTC	XhoI
771	Forward	CGCGGATCCCATATG-TCCGTATATCGCACCTTC	BamHI-
	Reverse	CCCGCTCGAG-CGGTTCTTTAGGTTTGAG	NdeI
772	Forward	CGCGGATCCCATATG-TTGCGGCGTTGGTGG	XhoI
	Reverse	CCCGCTCGAG-CAATGCCGACATCAAACG	BamHI-
774	Forward	CGCGGATCCCATATG-TCCGTTTACCCGTTCC	NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
790	Forward	CGCGGATCCCATATG-GCAAGAAGGTCAAAAAC	BamHI-
			NdeI

	Reverse	CCCGCTCGAG-GGCGTTGTTCCGATTTCG	XhoI
900	Forward	CGCGGATCCCATATG-CCGTCTGAAATGCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC	XhoI
901	Forward	CGCGGATCCCATATG-CCCGATTTTTCGATG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAAATGGAACAATACCAGG	XhoI
902	Forward	CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC	EcoRI- NdeI
	Reverse	CCCGCTCGAG-AAAAATGTACAATGGCGTAC	XhoI
903	Forward	CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT	EcoRI- NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA	XhoI
904	Forward	AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC	Kpn I
	Reverse	AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
904a	Forward	AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA	Eco RI
	Reverse	AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
905	Forward	CGCGGATCCCATATG-AACAAAATATACCGCATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCACTGATAACCGACAGAT	XhoI
907	Forward	CGCGGATCCCATATG-GGCGCGCAACGTGAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGCCACTGCCAGCG	XhoI
908	Forward	AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC	Eco RI
	Reverse	AAACTGCAG-TTAATATGGTTTTGTCGTTTCG	Pst I
909	Forward	CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT	XhoI
910	Forward	AAAGAATTC-GCATTTGCCGGCGACTCTGCCGAGCG	Eco RI
	Reverse	AAACTGCAG-TCAGCGATCGAGCTGCTCTTT	Pst I
911	Forward	AAAGAATTC-GCTTCCGCGTGGCCGGCGGTGC	Eco RI
	Reverse	AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTCCGC	Pst I
912	Forward	AAAAAAGAATTC- CAAATCCGTCAAACGCCACTCAAGTATTGAG	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC	Pst I
913	Forward	CGCGGATCCCATATG-GAAACCCGCCCCGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGTTGTGTTCCAGGTTG	XhoI
915	Forward	CGCGGATCCCATATG-TGCCGGCAGGCGGAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAAAATATAGGTATCAGG	XhoI
914	Forward	AAAGAATTC-GACAGAATCGGCGATTGGAAGCACG	Eco RI
	Reverse	AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGG	Pst I
916	Forward	CGCGGATCCCATATG-GCAATGATGGCGGCTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI

917	Forward	AAAAAAGAATTC-CCTGCCGAAAAACCGGCACCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCCCCGCCTTCACATCCTG	Pst I
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG	NdeI
			XhoI
920	Forward	CGCGGATCCCATATG-CACCGCGTCTGGGTC	BamHI-
	Reverse	CCCGCTCGAG-ATGGTGCGAATGACCGA	NdeI
			XhoI
921	Forward	AAAAAAGAATTC-TTGACGGAAATCCCCGTGAATCC	Eco RI
	Reverse	AAAAAACTGCAG-TCATTTCAAGGGCTGCATCTTCAT	Pst I
922	Forward	CGCGGATCCGCTAGC-TGTACGGCGATGGAGGC	BamHI-
	Reverse	CCCGCTCGAG-CAATCCCGGGCCGCC	NheI
			XhoI
923	Forward	CGCGGATCCCATATG-TGTTACGCAATATTGTCCC	BamHI-
	Reverse	CCCGCTCGAG-GGACAAGGCGACGAAG	NheI
			XhoI
925	Forward	CGCGGATCCCATATG-AAACAAATGCTTTTAGCCG	BamHI-
	Reverse	CCCGCTCGAG-GCCGTTGCATTTGATTTC	NdeI
			XhoI
926	Forward	CGCGGATCCCATATG-TGCGCGCAATTACCTC	BamHI-
	Reverse	CCCGCTCGAG-TCTCGTGCGCGCCG	NdeI
			XhoI
927	Forward	CGCGGATCCCATATG-TGCAGCCCCGCAGC	BamHI-
	Reverse	CCCGCTCGAG-GTTTTTTGCTGACGTAGT	NdeI
			XhoI
929a	Forward	AAAAAAGAATTC-CGCGGTTTGCTCAAAACAGGGCTGGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAAGAAAGACGGAACTACTGCC	Xba I
931	Forward	AAAAAAGAATTC-GCAACCCATGTTTTGATGGAAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCCCCGACAACAACGCGACG	Pst I
935	Forward	AAAAAAGAATTC-GCGGATGCGCCCGGATTTTGGATGACAAGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAAACCGCCAATCCGCCGACAC	Pst I
936	Forward	CGCGGATCCCATATG-GCCGCCGTCGGCGC	BamHI-
	Reverse	CCCGCTCGAG-GCGTTGGACGTAGTTTTG	NdeI
			XhoI
937	Forward	AAAAAAGAATTC-CCGGTTTACATTCAAACCGGCGCAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAATGTATGCTGTACGCCAAA	Pst I
939a	Forward	AAAAAAGAATTC-GGTTCCGGCAGCTGTGATGAAACC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGCAAACCTTGGATAAAGTTGGC	Pst I
950	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI-
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	NdeI
			XhoI
953	Forward	CGCGGATCCCATATG-GCCACCTACAAAGTGGAC	BamHI-
	Reverse	CCCGCTCGAG-TTGTTTGGCTGCCTCGAT	NdeI
			XhoI
957	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI-

	Reverse	CCCGCTCGAG-GTGAGAAAGGTCGCGC	NdeI
958	Forward	CGCGGATCCCATATG-GCCGATGCCGTTGCG	XhoI
			BamHI-
			NdeI
	Reverse	GCCCAAGCTT-GGGTCGTTTGTTCGTC	HindIII
959	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GTCGTCGCGGCGGGA	XhoI
961	Forward	CGCGGATCCCATATG-GCCACAAGCGACGACG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CCACTCGTAATTGACGC	XhoI
972	Forward	AAAAAAGAATTC- TTGACTAACAGGGGGGGAGCGAAATTAAAAAC	EcoRI
	Reverse	AAAAAATCTAGA-TTAAAAATAATCATAATCTACATTTTG	XbaI
973	Forward	AAAAAAGAATTC-ATGGACGGCGCACACCAGAAAAC	EcoRI
	Reverse	AAAAAACTGCAG-TTACTTCACGCGGGTCCCATCAGCGT	PstI
982	Forward	CGCGGATCCCATATG-GCAGCAAAGACGTAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CATCATGCCGCCCATCC	XhoI
983	Forward	CGCGGATCCCATATG-TTAGCTGTTGCAACAACAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
987	Forward	CGCGGATCCCATATG-CCCCACTGGAAGAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TAATAAACCTTCTATGGGC	XhoI
988	Forward	CGCGGATCCCATATG-TCTTTAAATTTACGGGAAAAAG	BamHI-
			NdeI
	Reverse	GCCCAAGCTT-TGATTTGCCTTTCCGTTTT	HindIII
989	Forward	CCGGAATTCTACATATG-GTCCACGCATCCGGCTA	EcoRI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGAATTTGTAGGTGTATTGC	XhoI
990	Forward	CGCGGATCCGCTAGC-TTCAGAGCTCAGCTT	BamHI-
	2		NheI
	Reverse	CCCGCTCGAG-AAACAGCCATTGTAGCGA	XhoI
992	Forward	CGCGGATCCCATATG-GACGCGCCCGCCCG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CCAAATGCCCAACCATTC	XhoI
993	Forward	CGCGGATCCCATATG-GCAATGCTGATTGAAATCA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GAACACATCGCGCCCG	XhoI
996	Forward	CGCGGATCCCATATG-TGCGGCAGAAAATCCGC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCTAAACCCCTGTTTTCTC	XhoI
997	Forward	CCGGAATTCTAGCTAGC-CGGCACGCCGACGTT	EcoRI-
			NheI
	Reverse	CCCGCTCGAG-GACGGCATCGCTCAGG	XhoI

Underlined sequences indicate restriction recognition sites.

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 1>:

```
g001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG GTGTCGGCGA ACGAGGTGTC
51  CGGCAGGGCT TCGCGCCGGA TGGTGCTGGT CATCTGCCAG ACGCTGCCGA
101 AACGCGATAC TTTAAACGGC TCGGGTACGC ATACTTTACC GGTTCGGCG
151 ATTTGCCGA GGTGTTGCG CAGCAAATCG ACAATCATCA CGTTTCGGC
201 GCGGTTTTTC GGGTCGGTTT GTAATCGGC GCGCGGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTGGT GCTGATGGTG
301 CCGTCTGAAG CGATGTTGAG GAAGAGTCG GCGGAGAAC ACAGCGTCCA
351 CCGGATTGC CCGCTTCAT CGGCAGGTG GGACAATACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 2; ORF 001.ng>:

```
g001.pep
1  MLPOGKAARR VSANEVSGRA CARMVLVICQ TLPKRDTLNG SGHTLTPVWA
51  ILPRSLRSKS TIITFSARFF GSVNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEAMLRKSS GEKHSVHAD C PASSGRWDNT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3>:

```
m001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTC
51  CGGCAssCTT ss.GCTTGGA yGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTCGGCG
151 ATTTGCCGA GATCGTTACG CAGCAAATCG ACAATCATCA CGTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GCGCGGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTGGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTCG GCGGAGAAC ACAGCGTCCA
351 CCGGATTGC CCCTCCGCAT CGGCAGGTG GGACAAGACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 001>:

```
m001.pep
1  MLPOGKAARR MSANEVCGXL XAWXVLVICQ TLPKRDTLNG SGHTVTPVWA
51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEPILRKSS GEKHSVHAD C PSASGRWDKT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 5>:

```
a001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTC
51  CGGCAAGGCT TGGGCTTGGA TGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTCGGCG
151 ATTTGCCGA GGTGTTACG CAGCAAATCG ACAATCATCA CGTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GCGCGGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTGGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTCG GCGGAGAAC ACAGCGTCCA
351 CCGGATTGC CCTTGTGCAT CGGCAGGTG GGACAAAACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 6; ORF 001.a>:

```
a001.pep
1  MLPOGKAARR MSANEVCGKA WAWMVLVICQ TLPKRDTLNG SGHTVTPVWA
51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
101 PSEPILRKSS GEKHSVHAD C CASGRWDKT A*
```

m001/a001 96.2% identity over a 131 aa overlap

	10	20	30	40	50	60
m001.pep	MLPOGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGHTVTPVWAILPRSLRSKS					
a001.pep	MLPOGKAARRMSANEVCGKAWAWMVLVICQTLPKRDTLNGSGHTVTPVWAILPRSLRSKS					

166

	10	20	30	40	50	60
m001.pep	70	80	90	100	110	120
	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
a001.pep						
	70	80	90	100	110	120
	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
	130					
m001.pep	PSASGRWDKTAX					
a001.pep	PCASGRWDKTAX					
	130					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 001 shows 89.3% identity over a 131 aa overlap with predicted ORF (ORF 001.ng) from *N. gonorrhoeae*:

**m001/g001**

	10	20	30	40	50	60
m001.pep	MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS					
g001	MLPQGKAARRVSAANEVSGRACARMVLVICQTLPKRDTLNGSGTHTLPVWAILPRSLRSKS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
g001	TIIITFSARFFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHADC					
	70	80	90	100	110	120
	TIIITFSARFFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHADC					
	130					
m001.pep	PSASGRWDKTAX					
	::					
g001	PASSGRWDNTAX					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 7>:

**g003.seq**

```

1  ATGGTCGTAT TCGTGGCTGA AGGCGTATTC GGTGCGCTG TTTTGGGTCA
51  CTGGGTATTG CTCTTCGETC AGGGTGCGTT TGAGTTCGGC GTCACCTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGCTTTGGT
151 TTGCCCCGGC AGCGGTTCGT CGGCTTTGCG GATGTCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAGG TAGTCCTGAT GGTATTCCTC GGCCTCGTAG
251 AAGTTTTTCA GCGGTTTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTC
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
351 TGAGTACAC GCCGCTGCGG TATTGCGTGC CGGTGTCGTT ACCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAA TAATATTGCA GGATGTCGTC
451 CAGgCTGagt TTGTCGGCAT CGTaggtcac tTTGACGGTC TCGGCATGAC
501 CCGTATGGCG GTaggacact tctTCgtanc TcGGGtTTTC CGTGtGCCG
551 TTGGCgttac cGGATACCGC gtcaACCACG CCGTcgatgc gttggaAATa
601 ggCTTCCAAG ccccaaaagc agccgccggc gaagtaaagt gtgcccgtgt
651 tcatgattGC TGA

```

This corresponds to the amino acid sequence <SEQ ID 8; ORF 003.ng>:

**g003.pep**

```

1  MVVFVAEGVF GRAVLGHLVL LFGQGAFFFG VTRFFIRCRV EAFALRCGFG
51  FARQRFVGFa DVDVAVAVGV FNQVVLMLVL GVVEVFQRFV FNNEGQLVFL
101 LLAFEGGGDD GFFGGVGvvh AAVLRAGVV TLFVEAGRIN DAEIILQDVV

```

151 QAEFVGIVGH FDGLGMTRMA VGHFFVRVFR VAVGVTGYRV NHAVDALZIG  
201 FQAPKAAAGE VNGARVHDC

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 9>:

m003.seq  
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCCGCTG TTTTGGGTAA  
51 CTTGCTATTG CTCTTCGGTC AGGGTGCCTT TGAGTTCGGC GTCACCTCGT  
101 TTTTATACG TTGCCCGCTC GAAGCCTTTG CCTTGCCTGG CCGTCTTGGT  
151 TTTGCCCGGC AGCGGTTCGT CAGCkTTGCG GATGTCGATG TGGCAGTAGC  
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG  
251 AAGTTTTtCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTGG  
301 CTGCTCGCGT TTGAGGGCGk CGGCGATGAC GGCTTTTTTCG kCGGGGTCGG  
351 TGTAGTACAC GCCGCTGCGG TATTGCGTAC CCGTGTCTGT GCCCTGTTTG  
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC  
451 TAGGCTGAGT TTGTGCGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC  
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTGGCCG  
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAAATA  
601 GGCTTCCAAG CCCAGAAGC AGCg.CCGGC GAGGTAAATG GTGCGCGTGT  
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 10; ORF 003>:

m003.pep Length: 221  
1 MVVFVAEGIF GRAVLGNLXL LFGQGAFEFV VTRFFIRCRV EAFALRGGLG  
51 FARQRFVSXA DVDVAVAVGV FNQVLMVFL GIVEVFQRLV FNNEGQLVFL  
101 LLAFEGXGDD GFFXGVGVVH AAAVLRTGVV ALFVEAGRIN DAEEILQDVV  
151 \*AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDALZIG  
201 GFQAPKAAAGE EVNGARVHDC \*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 11>:

a003.seq  
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCCGCTG TTTTGGGTAA  
51 CTTGCTATTG CTCTTCGGTC AGGGTGCCTT TGAGTTCGGC GTCACCTCGT  
101 TTTTATACG TTGCCCGCTC GAAGCCTTTG CCTTGCCTGG CCGTCTTGGT  
151 TTGCCCCGC AGCGGTTCGT CCGCTTTGCG GATATCGATG TGGCAGTAGC  
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG  
251 AAGTTTTTCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTGG  
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTTCG GCGGGGTCGG  
351 TGTAGTACAC GCCGCTGCGG TATTGCGTAC CCGTGTCTGT GCCCTGTTTG  
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC  
451 TAGGCTGAGT TTGTGCGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC  
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTGGCCG  
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAAATA  
601 GGCTTCCAAG CCCAGAAGC AGCCGCGGCG GAGGTAGATG GTGCGCGTGT  
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 12; ORF 003.a>:

a003.pep  
1 MVVFVAEGIF GRAVLGNLVL LFGQGAFEFV VTRFFIRCRV EAFALRCGLG  
51 FARQRFVGFA DIDVAVAVGV FNQVLMVFL GIVEVFQRLV FNNEGQLVFL  
101 LLAFEGGGDD GFFGGGVGVH AAAVLRTGVV ALFVEAGRIN DAEEILQDVV  
151 \*AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDALZIG  
201 GFQAPKAAAGE EVDGARVHDC \*

m003/a003 95.9% identity over a 220 aa overlap

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVLGNLXL	LFGQGAFEFV	VTRFFIRCRV	EAFALRGGLG	FARQRFVSXA	
						:
a003	MVVFVAEGIFGRAVLGNLVL	LFGQGAFEFV	VTRFFIRCRV	EAFALRCGLG	FARQRFVGFA	
	10	20	30	40	50	60

168

	70	80	90	100	110	120
m003.pep	DVDVAVAVGVFNQVVLVFLGIVEVFQRLVFNNQQLVFLLLAFEGXGDDGFFXGVGVVH					
a003	DIDVAVAVGVFNQVVLVFLGIVEVFQRLVFNNQQLVFLLLAFEGGDDGFFGGVGVVH					
	70	80	90	100	110	120
m003.pep	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMVGHVFIARIF					
a003	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMVGHVFIARIF					
	130	140	150	160	170	180
m003.pep	RVAVGVAGYRVNHAVDALFQAPAAAXGEVNGARVHDFX					
a003	RVAVGVAGYRVNHAVDALFQAPAAAGEVDGARVHDFX					
	190	200	210	220		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 003 shows 88.6% identity over a 219 aa overlap with a predicted ORF (ORF 003.ng) from *N. gonorrhoeae*:

m003/g003

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVLGNLXLLFGQGAFFGVTRFFIRCRVEAFALRGGLGFARQRFVSXA					
g003	MVVFVAEGVFGRAVLGHLVLLFGQGAFFGVTRFFIRCRVEAFALRCGFGFARQRFVGF					
	10	20	30	40	50	60
m003.pep	DVDVAVAVGVFNQVVLVFLGIVEVFQRLVFNNQQLVFLLLAFEGXGDDGFFXGVGVVH					
g003	DVDVAVAVGVFNQVVLVFLGVVEVFQRFVNNQQLVFLLLAFEGGDDGFFGGVGVVH					
	70	80	90	100	110	120
m003.pep	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMVGHVFIARIF					
g003	AAAVLRAGVVTLFVEAGRINDAEEILQDVVQAEFVGIVGHFDGLGMRMAVGHFFV-RVF					
	130	140	150	160	170	180
m003.pep	RVAVGVAGYRVNHAVDALFQAPAAAXGEVNGARVHDFX					
g003	RVAVGVAGYRVNHAVDALFQAPKAAAGEVNGARVHDC					
	190	200	210	220		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 13>:

g004.seq

```

1  ATGgtagAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTGAT
51  GCGCCCATGC CAACAagtga gccAAatgtT CGGCGGCAGG GCCTacgatT
101 TCCGCGCCGA TAAagcggcc gGTGgctTTT tcgGCataca ggcgcaTatg
151 gCCTTTGTTT ACCAgcatca cgcggtcg accttgaTTT TTGAACGATA
201 CTTCCGCCgaT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTCAAAC CGACAAAGCC GATTTGCgga ctggtaaACA CCACGCCAAT
301 GGTgctgcgg cGCAAACCGC TGCCGATAtt cgGtagcgg ccccgcgtaa
351 ttgcccggca atcttacctt ggtcggcggc ttcatGCAGC AGGGGCagtt
401 gggttgacgc gtcgcccgcg ataAAGATAT GCGGAATgct ggtCTGCATg
451 gtCAGCGGAT CGGCAACGGG tacgcccgc gcgtctttgT CGATATTGAT
501 GTTTTCCAAA CCGATAtgt CAACGTTCGG ACGGCgACCT ACGGCTGCCA

```



```

551 ACATATATTC GGCAACAAA. ACGCCTTTTT CGCCATCCTG CTCCCAATGG
601 ACTtctACAT TGCCGTCTGC GTCGAGTTTG ACCTCGGTTT TAGCATCCAG
651 ATGCAGTTTC AATtctTCTC CGAACACGGC TTTCCGCTCG TCTGAAACAA
701 CGGGGTGCGA AATGCCGCCG ATGATTCCGC CCAAACCGAA AATTTCAACT
751 TTCACACCCA AACGGTGCAA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 14; ORF 004.ng>:

**g004.pep**

```

1 MVERHIQHLR NGHLLHMRPC QQVSQMFQGR AYDFRADKAA GGFFGIQAHM
51 AFVYQHHAHA TLIFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAN
101 GAAAQTAADI RVAAPRYCPA ILPWSAASCS RGSWLDASPA IKICGMLVCM
151 VSGSATGTPR ASLSILMFSK PILSTFGRRP TAANIYSATN TPFSPSCSQW
201 TSTLPSASSL TSVLASRCSF NSSPNTAFAS SETTGSEMPP MIPPKPKIST
251 FTPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 15>:

**m004.seq**

```

1 ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
51 GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCCAGG GCCTACGATT
101 TCCGCGCCGA TAAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGTTT TTGAACGATA
201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTAGAC CGACAAAGCC GATTTGCGGA CTGGTAAACA CCACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGA GGCAGTTGGT
401 TGGACGCATC GCCTGCGATG AAGATATGCG GAATACTGGT CTGCATGGTC
451 AGCGGGTCGG CAACAGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATATT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCATC GAGTTTGACC TCGGTTTTAG CATCCAGATG
651 CAGTTTCAAT TCTTCGCCGA ACACGGCGTT CGCCTCGTCT GAAACGACGG
701 GGTGCGAAAT GCCGCCGATG ATTCCGCCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCAAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 16; ORF 004>:

**m004.pep**

```

1 MVERHIQHLR NGHLLHMCPS QQVRQMFQGR AYDFRADKAA GGFFGIQAHM
51 AFVHQHHAHA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAD
101 GAAPQTAADI RVAAALSPA ILPWSAASCSR GSWLDASPA KICGILVCMV
151 SGSATGTPRA SFSILIFSKP ILSTFGRRPT AASIYSATNT PFSPPSCSQW
201 STLPSASSLT SVLASRCSFN SSPNTAFASS ETTGSEMPPM IPPKPKISTF
251 TPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 17>:

**a004.seq**

```

1 ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT
51 GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCCAGG ACCTACGATT
101 TCTGCGCCGA TGAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGTTT TTGAACGATA
201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTCAAAC CGACAAAGCC GATTTGCGGA CTGGTGAACA CTACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGGCAATC TTGCCTTGGT CGGCGGCTTC ATGCAGCAGG GGCAGTTGGT
401 TGGACGCGTC GCCCGCAATA AAGATATGCG GAATACTGGT CTGCATAGTC
451 AGCGGATCGG CAACGGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATGTT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCTACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCGTC GAGTTTGACC TCGGTTTTAG CATCCAAATG
651 CAGTTTCAAT TCTTCACCGA ACACGGCTTT CGCCTCGTCT GAAACGACGG
701 GGTGCGAAAT GCCGCCGATG ATGCCACCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCAAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 18; ORF 004.a>:

**a004 . pep**

```

1  MVERHIQHLR NGHLLHMCPS QQVRQMFGR TYDFCADEAA GGFFGIQAHM
51  AFVYQHHAHA ALVFERFYFAD DKFVGLVLRG NLRVFQTDKA DLRTGEHYAD
101 GAAQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAI KICGILVCIV
151 SGSATGTpra SFSILMFskP ILSTFGRRPT AASIYSATNT PFSPSCSQWT
201 STLPSASSLA SVLASKCSFN SSPNTAFASS ETTGSEMPPM MPPKPKISTF
251 TPKRCNA*
```

**m004/a004** 94.9% identity over a 257 aa overlap

m004 . pep	10	20	30	40	50	60
	MVERHIQHLRNGHLHLMCPSQQVRQMFGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA					
a004	MVERHIQHLRNGHLHLMCPSQQVRQMFGRRTYDFCADEAAGGFFGIQAHMAFVYQHHAHA					
	10	20	30	40	50	60
m004 . pep	70	80	90	100	110	120
	ALVFERFYADDKFVGLVLRGNLRFQTDKADLRTGKHHADGAAPQTAADIRVAAALSPAI					
a004	ALVFERFYADDKFVGLVLRGNLRFQTDKADLRTGEHYADGAAQTAADIRVAAALSPAI					
	70	80	90	100	110	120
m004 . pep	130	140	150	160	170	180
	LPWSAASCSRGSWLDASPAIKICGILVCMVSGSATGTPRASFILFSKPISTFGRRPT					
a004	LPWSAASCSRGSWLDASPAIKICGILVCIVSGSATGTPRASFILMFskPILSTFGRRPT					
	130	140	150	160	170	180
m004 . pep	190	200	210	220	230	240
	AASIYSATNTPFSPSCSQWTSTLPSASSLTVLASRCSFNSSPNTAFASSETTGSEMPPM					
a004	AASIYSATNTPFSPSCSQWTSTLPSASSLASVLASKCSFNSSPNTAFASSETTGSEMPPM					
	190	200	210	220	230	240
m004 . pep	250					
	IPPKPKISTFTPKRCNAX					
a004	MPPKPKISTFTPKRCNAX					
	250					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 004 shows 93.4% identity over a 258 aa overlap with a predicted ORF (ORF 004.ng) from *N. gonorrhoeae*:

**m004/g004**

m004 . pep	10	20	30	40	50	60
	MVERHIQHLRNGHLHLMCPSQQVRQMFGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA					
g004	MVERHIQHLRNGHLHLMRCPQQVQMFGRAYDFRADKAAGGFFGIQAHMAFVYQHHAHA					
	10	20	30	40	50	60
m004 . pep	70	80	90	100	110	119
	ALVFERFYADDKFVGLVLRGNLRFQTDKADLRTGKHHADGAAPQTAADIRVAAA-LSPA					
g004	TLIFERYFADDKFVGLVLRGNLRFQTDKADLRTGKHHANGAAAQTAADIRVAAAPRYCPA					
	70	80	90	100	110	120
m004 . pep	120	130	140	150	160	179
	ILPWSAASCSRGSWLDASPAIKICGILVCMVSGSATGTPRASFILFSKPISTFGRRP					

```

|||||
g004      130      140      150      160      170      180
          180      190      200      210      220      230      239
m004.pep  TAASIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMP
          |||||
g004      TAANIYSATNTPFSPSCSQWTSTLPSASSLTSVLASRCSFNSSPNTAFASSETTGSEMP
          190      200      210      220      230      240

          240      250
m004.pep  MIPPKPKISTFTPKRCNAX
          |||||
g004      MIPPKPKISTFTPKRCNA
          250

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 19>:

```

g005.seq
1  ATGGGGATGG ACAATATTGA TATGTTTCATG CCTGAACAAG AGGAAATCCA
51  ATCAATGTGG AAAGAAATTT TACTGAATTA CGGTATTTTC CTGCTCGAAC
101 TGCTTACCGT GTTCGGCGCA ATTGCGCTGA TTGTGTTGGC TATCGTACAG
151 AGTAAGAAAC AGTCGGAAAG CGGCAGTGTC GTACTGACAG ATTTTTCGGA
201 AAATTATAAA AAACAGCGGC AATCGTTTGA AACATTCTTT TTAAGCGAGG
251 AAGAGACAAA ACATCAGGAA AAAAAAGAAA AGAAAAAGGA AAAGGCGGAA
301 GCCAAAGCAG AGAAAAAGCG TTTGAAGGAG GCGGGGGAGA AATCTGCCGA
351 AACGCAAAAA TCCCGCCTTT TTGTGTTGGA TTTTGACGGC GATTTGTATG
401 CACACGCCGT AGAATCCTTG CGTCATGAGA TTACGGCGGT GCTTTTGATT
451 GCCAAGCCTG AAGATGAGGT TCTGCTCAGA TTGGAAAGTC CGGGCGCGCT
501 GGTTACCGGT TACGTTTGG CGGCTTCGCA GCTTAGGCGT TTGCGCGAAC
551 GCAATATTCC GCTGAccgtc gccgTCGATA AGGTCGCGGC AAGCGgcggc
601 tatatgatgg cgtgtgtgGC GGATAAAAT GTTCCGCTc cgtttgcggt
651 catcggttcg gtgggtgtgg tgGcggaagt gcCGAATATC CaccgCctGT
701 TGAAAAACA TGATATTGAT GTGGATGTGA TGACGGCGGG CGAATTTAAG
751 CGCACGGTTA CTTTTATGGG TGAAATACG GAAAAGGGCA AACAGAAATT
801 CCGGCAGGAA CTGGAGGAAA CGCATCAGTT GTTCAAGCAG TTTGTCACTG
851 AAAACCGCCC CGGGTTGGAT ATTGAAAAA TAGCGACGGG CGAGCATTGG
901 TTCGCCGGC AGGCGTTGGC GTTGAAGTGT ATTGACGAGA TTTCGACCAG
951 TGATGATTG TTGTTGAAAG CGTTGAAAA CAAACAGGtt aTCGAAGTGA
1001 AATATCAGGA GAAGCGAAGC CTGATCCAGC GCATTGGTTT GCAGGCGGAA
1051 GCTTCCGTTG AAAAGTTGTT TGCCAACTT GTCAACCGGC GAGCGGATGT
1101 GATGTAG

```

This corresponds to the amino acid sequence <SEQ ID 20; ORF 005.ng>:

```

g005.pep
1  MGMNDNIDFM PEQEEIQSMW KEILLNYGIF LLELLTVFGA IALIVLAIVQ
51  SKKQSESGSV VLTDFSENYK KQRQSFETFF LSEEETKHQE KKEKKKEKAE
101 AKAEEKRLKE GGEKSAETQK SRLFVLDFDG DLYAHAVESL RHEITAVLLI
151 AKPEDEVLLR LESP GG VVHG YGLAASQLRR LRERNIPLTV AVDKVAASGG
201 YMMACVADKI VSAPFAVIGS VGVVAEVPNI HRLLKKHDID VDVMTAGEFK
251 RTVTFMGENT EKGKQKFRQE LEETHQLFKQ FVSENRPGLD IEKIATGEHW
301 FGRQALALNL IDEISTSDDL LLKAFENKQV IEVKYQEKRS LIQRIGLQAE
351 ASVEKLFACL VNRRADVM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 21>:

```

m005.seq
1  ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51  GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AAACAGTCGG AwAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
201 TAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CACAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA

```

```

301 GCAGAGAAAA A.CGTTTGAA GGAGGGTGGG GAGAAATCTG CCGAAACGCA
351 nAAATCACGC CTTTTGTGT TGGANNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGCGAGCGG CGGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGCTTCC GCTCCGTTTG CGATTGTCGG
651 TTCGGTGGGT GTGGTGGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAAACAGA AATTCCGACA
801 GGAACGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTATATCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTCAGGC GGAAGCTTCT
1051 GTTGAAAAGT TGTTTGCCAA ACTTGTC AAC CGGCGGGCGG ATGTGATGT A
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 22; ORF 005>:

**m005.pep**

```

1 MDNIDMFMP EEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQXSGSVVL TDFSENYKKQ RQSFEEFFLS GEEA... EKKKEKAEAK
101 AEKXRLKEGG EKSAETXKSR LFVLXXXXXX XXXXXXXXXX XXXXXXXXXX
151 XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXASGGYM
201 MACVADKIAS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG
301 RQALALNLID EISTSDLLL KAFENKQVIE VKYQEKQSLI QRIGLQAEAS
351 VEKLFKLVN RRADVM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 23>:

**a005.seq**

```

1 ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51 GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AAACAGTCGG AAAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
201 TAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CAAAACATCA GAAAAAGAG GAAAAGAAA AGGAAAAGGC GGAAGCCAAA
301 GCAGAGAAAA AGCGTTTGAA GGAGGGTGGG GAGAAATCTT CCGAAACGCA
351 AAAATCCCGC CTTTTGTGT TGGATTTGA CGGCGATTTG TATGCACACG
401 CCGTAGAATC CTTGCGTCAT GAGATTACGG CGGTGCTTTT GATTGCCAAG
451 CCTGAAGATG AGGTTCTGCT TAGATTGGAA AGTCCGGGCG GCGTGGTTCA
501 CGGTACGGT TTGGCGGCTT CGCAGCTTAG GCGTTTGC GCACGCAATA
551 TTCCGCTGAC CGTCGCCGTC GATAAGGTGG CGGCGAGCGG TGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGTTTCC GCTCCGTTTG CGATTGTCGG
651 TTCGGTGGGT GTTGTAGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAAACAGA AATCCGACA
801 GGAACGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTATATCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTCAGGC GGAAGCTTCT
1051 GTTGAAAAGT TGTTTGCCAA ACTTGTC AAC CGGCGGGCGG ATGTGATGTA
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 24; ORF 005.a>:

**a005.pep**

```

1 MDNIDMFMP EEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQESGSGSVL TDFSENYKKQ RQSFEEFFLS GEEAKHQEKE EKKKEKAEAK
101 AEKXRLKEGG EKSETQKSR LFVLDLFDGL YAHAVESLRH EITAVLLIAK
151 PEDEVLLRLE SPGGVVHGYG LAASQLRLRL ERNIPLTVAV DKVAASGGYM
201 MACVADKIVS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG

```

301 RQALALNLID EISTSDDLL KAFENKQVIE VKYQEKQSLI QRIGLQAEAS  
 351 VEKLFKLVN RRADVM\*

m005/a005 79.2% identity over a 366 aa overlap

	10	20	30	40	50	60
m005.pep	MDNIDMFMPQE	EIQSMWKEILL	NYGIFLLELL	TVFGAIALIV	LAIVQSKKQ	SXSGSVVL
a005	MDNIDMFMPQE	EIQSMWKEILL	NYGIFLLELL	TVFGAIALIV	LAIVQSKKQ	SESGSVVL
	10	20	30	40	50	60
m005.pep	70	80	90	100	110	120
	TDFSENYKKQ	RQSFEAFFLS	GEEAQHQEKE	KKKKEKAEAK	AEKKRLKEG	GEKSAETXKSR
a005	TDFSENYKKQ	RQSFEAFFLS	GEEAKHQEKE	KKKKEKAEAK	AEKKRLKEG	GEKSSSETQKSR
	70	80	90	100	110	120
m005.pep	130	140	150	160	170	180
	LFVLXXXXXXXX	XXXXXXXXXXXX	XXXXXXXXXXXX	XXXXXXXXXXXX	XXXXXXXXXXXX	XXXXXXXXXXXX
		:	:	:	:	:
a005	LFVLDFDGDLY	AHAVESLRHEI	TAVLLIAKPE	DEVLLRLESP	GGVVHGYGLA	ASQLRRLR
	130	140	150	160	170	180
m005.pep	190	200	210	220	230	240
	XXXXXXXXXXXX	ASGGYMMACV	ADKIASAPFA	IVGSVGVVAE	VPNIHRLKKH	IDIDVD
	:					
a005	ERNIPLTVAV	DKVAASGGYMM	ACVADKIVSA	PFAIVGSVGV	VAEVPNIHRL	KKHIDIDVD
	190	200	210	220	230	240
m005.pep	250	260	270	280	290	300
	VMTAGEFKRT	VTFMGENTEK	GKQKFRQELE	ETHQLFKQFV	SENRPQLDIE	EVATGEHWFG
a005	VMTAGEFKRT	VTFMGENTEK	GKQKFRQELE	ETHQLFKQFV	SENRPQLDIE	EVATGEHWFG
	250	260	270	280	290	300
m005.pep	310	320	330	340	350	360
	RQALALNLID	EISTSDLLKAF	ENKQVIEVKY	QEKQSLIQRI	GLQAEASVEK	LFAKLVN
a005	RQALALNLID	EISTSDLLKAF	ENKQVIEVKY	QEKQSLIQRI	GLQAEASVEK	LFAKLVN
	310	320	330	340	350	360
m005.pep	RRADVMX					
a005	RRADVMX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 005 shows 77.0% identity over a 366 aa overlap with a predicted ORF (ORF 005.ng) from *N. gonorrhoeae*:

m005/g005

	10	20	30	40	50
m005.pep	MDNIDMFMPQE	EIQSMWKEILL	NYGIFLLELL	TVFGAIALIV	LAIVQSKKQ
g005	MGMDNIDMFMPQE	EIQSMWKEILL	NYGIFLLELL	TVFGAIALIV	LAIVQSKKQ
	10	20	30	40	50
m005.pep	60	70	80	90	100
	VLTDSENYKKQ	RQSFEAFFLS	GEEAQHQEKE	KKKKEKAEAK	AEKKRLKEG
g005	VLTDSENYKKQ	RQSFEAFFLS	GEEAQHQEKE	KKKKEKAEAK	AEKKRLKEG
	70	80	90	100	110
	120				

	120	130	140	150	160	170
m005.pep	SRLFVLXXX					
g005	SRLFVLDGDLIAHVESLRHEITAVLLIAKPEDEVLLRLESPPGGVVHGYGLAASQLRR					
	130	140	150	160	170	180
	180	190	200	210	220	230
m005.pep	XXXXXXXXXXXXXXXXXASGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLKKHDID					
	:      :      :      :      :      :					
g005	LRERNIPLTVAVDKVAASGYMMACVADKIVSAPFAVIGSVGVVAEVPNIHRLKKHDID					
	190	200	210	220	230	240
	240	250	260	270	280	290
m005.pep	VDVMTAGEFKRTVTFMGENTTEKQKQKFRQLEETHQLFKQFVSENRPQLDIEEVATGEHW					
	:      :      :      :      :					
g005	VDVMTAGEFKRTVTFMGENTTEKQKQKFRQLEETHQLFKQFVSENRPGLDIEKIATGEHW					
	250	260	270	280	290	300
	300	310	320	330	340	350
m005.pep	FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFACL					
	:      :      :      :      :					
g005	FGRQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFACL					
	310	320	330	340	350	360
	360					
m005.pep	VNRRADVMX					
g005	VNRRADVMX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 25>:

**g006.seq**

```

1  ATGCTGCTGG TGCTggaatt ttggttCGGc gtGtCGGCGG TGGGCatact
51  tgCGTTGTTT TTATGGCttt TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCGG CCTGAACAAC AGCTTGGAAC gcgACAACCA CTTTATCCGA
151 AAAGGCGACG AGCGGCAGCT GTACCGCCAT TACGGACTGG TTTGCGCGCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TCGTTCGGCG
251 CGGCGATGGG TATTTTGTTT GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACGGCA GCGCGGGGCA TATTTATTCTG GTCGGCACTT ATCTGTGGAT
351 GTTTGCCATG AGTTTGGACG ATGTGCCGCG ATTGGTCGAA CAATATTCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGCGAACG GAACATCAAA
451 GCCGGAACCTT GA

```

This corresponds to the amino acid sequence <SEQ ID 26; ORF 006.ng>:

**g006.pep**

```

1  MLLVLEFWFG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
51  KGDERQLYRH YGLVSRLRVL ISNREAPGYL CVGAAMGILF GFAFVMTLK
101 GYGSAGHIYS VGTYLWMPAM SLDDVPRLVE QYSLNKDIGQ RIEWSERNIK
151 AGT*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 27>:

**m006.seq**

```

1  ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
51  TGCGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCGG CCTGAACAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA
151 AAAGGCGACC GGCGGCAGCT GTACCGCCAT TACGGACTGC TTGCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TCGTTCGGCA
251 CGGCGATGGG TATTTTGTTT GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACAGCA GCGCGGGGCA TGTCTATTCTG GTCGGCACTT ATCTGTGGAT

```

```

351 GTTGGCCATG AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGGAACG GAACATCAAA
451 GCCGGAACCTGA

```

This corresponds to the amino acid sequence <SEQ ID 28; ORF 006>:

```

m006.pep
1  MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERD�HFIR
51  KGDRRQLYRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101 GYSSAGHVYS VGTYLWMFAM SLDDVPRLVE QYSLKDIGQ RIEWSEKNIK
151 AGT*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 29>:

```

a006.seq
1  ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
51  TCGGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCCT CCTGAAGAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA
151 AAAGGCGACG AGCGGCAGCT GGACCGCCAT TACGGAATGC TTGCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TCGCTCGGCA
251 CCGCGATGGG TATTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT
351 GTTTGCCATA AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGGAACG GAACATCAAA
451 GCCGGAACCT GA

```

This corresponds to the amino acid sequence <SEQ ID 30; ORF 006.a>:

```

a006.pep
1  MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLKN SLERD�HFIR
51  KGDERQLDRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101 GYSSAGHVYS VGTYLWMFAI SLDDVPRLVE QYSLKDIGQ RIEWSEKNIK
151 AGT*

```

m006/a006 96.7% identity over a 153 aa overlap

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVS					
a006	MLLVLEFWVGVS					
	10	20	30	40	50	60
m006.pep	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAM					
a006	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAI					
	70	80	90	100	110	120
m006.pep	SLDDVPRLVEQYSLKDIGQRIEWSERNIKAGTX					
a006	SLDDVPRLVEQYSLKDIGQRIEWSERNIKAGTX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 006 shows 95.4% identity over a 153 aa overlap with a predicted ORF (ORF 006.ng) from *N. gonorrhoeae*:

m006/g006

```

m006.pep
1  MLLVLEFWVGVS VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERD�HFIR KGDRRQLYRH

```

```
g006-1.seq
1  ATGTGGAAAA  TGTGTAAACA  CATAGCCAAA  ACCCTTTCGA  AGCGATTGAT
51  TGGCACAATT  TCCCGGTCG  GACTGGAAAA  CCTTTTGATG  CTGGGGTGATC
101 CGGTGTTTGG  CGGCTGGGCG  ATTAATGCCG  TGATTGCGGG  GAGGGTGTGG
151 CAGGCGTTGC  TGTACGCTTT  GGTGTATATT  TTGATGTGGC  TGGTCGGTGC
201 GGCACGGCGG  ATTCCCGATA  CGCCGACGTT  TACGCGGATT  TATACCGAAA
251 TCGCCGTGCC  GGTGTGTTG  GACAACGGC  AGCGGCAAGT  CCGCGATTCA
301 CGCGTAAC TG  CACGGGTTGC  CCTGTCGCGT  GAATTTGTCA  GCTTTTTTGA
351 AGAACACCTG  CCGATTGCCG  CGACATCCGT  CGTATCCATA  TTCGGCGCGT
401 GCATCATGCT  GCTGGTGCTG  GAATTTTGGG  TCGGCGTGTC  GCGCTGTGGG
451 ATACTTTCGT  TGTTTTATG  CGTTTGGCA  CGTTTGGCG  CCATCAGCGA
501 AAACCTGTAT  TTCCGCCTGA  ACAACAGCTT  GGAACGCGAC  AACCACCTTA
551 TCCGAAAAGG  CGACGAGCGC  CAGCTGTACC  GCCATTACGG  ACTGGTTTCG
601 CGCCTCGGTG  TGCTGATTTC  CAACCGCGAA  GCCTTCGGCT  ATCTCTGCGT
651 CGCGCGCGCG  ATGGGTATT  TGTTCCGCTT  TGCTTTTGTG  ATGATGACGC
701 TCAAAGGCTA  CGGCAGCGCG  GGGCATATTT  ATTGGTTCGG  CACTTATCTG
751 TGGATGTTTG  CCATGAGTTT  GGACGATGTG  CCGCGATTGG  TCGAACATAA
801 TTCCAATTTG  AAAGACATCG  GACAACGGAT  AGAGTGGTCG  GAACGGAACA
851 TCAAAGCCGG  AACTTGA
```

g006-1.pep

1	MWKMLKHI	AK	THRKRLIG	TF	SPVGLENLL	M	LGYPVFGG	WA	INAVIAGR	RVW
51	<u>QALL</u>	YALVVF	<u>LMWL</u>	VGAARR	IADTFTF	TRI	YTEIAVP	VPVL	EQQRQ	VPHS
101	AVTARVAL	SR	EFVSFF	EHL	PIAATSV	VSI	<u>FGACIM</u>	LLVL	EFWVG	SAVG
151	<u>ILAL</u>	FLWLLP	RFAAIS	ENLY	FRLNNS	LERD	NHFIRK	GDGR	QLYRHY	GLVS
201	RLRVLS	INRE	AFGYL	<u>CVGAA</u>	MGILEG	FAFV	MMTLKGY	GS	GHIYS	VTYL
251	WMFAMSL	DDV	PLRVEY	GSNL	KDIGOR	IEWS	ERNIKAG	T		

```
m006-1.seq
1  ATGTGGA AAA TGTGAAACA CATAGCCCCA ACCCACCGCA AGCGATTGAT
51  TGGCACA TTT TCCTGGTGC GACTGGA AAA CCTTTTGATG CTGGTGATC
101 CGGTGTTTG CGGCCGGCG GATCAATGCC TGATTGCGGG GGAGGTGTGG
151 CAGGCGT TGC TGTACGCTT GGTGTGCTT TTGATGTGCC TGGTCGGTGC
201 GGTGCGGCG GTTCCGATA CGCGCACGTT TACGCGGAT TATACGAAA
251 TCGCCGTGCC GCTCGGT TG AACACGCGC AGCGACAAGT CCGCGATTCG
301 GCGGTA ACTG CGCGGGTTG CCTGTCGCGT GAGTTTGTC GCTTTTTTGA
351 AGAACAC CTG CCGATTGCC GCAATCCGT CGTATCCATA TTCGGCGCGT
401 GCATCATG CT GCTGCTGCT GCAATTTGGG TCGCGGTGTC CGCGTGCGGC
451 ATACTT GCGT TGTTTTATG GCTTTTGCCA CGTTTGGCG CCGTCAGCGA
501 AAACCTG TAT TTCCGCCTA ACAACAGCTT GGAACAGCAG AACCACTTGA
551 TCCGAAA AGG CGACCGGCG CAGCTGTACC GCCATTCAGG ATCGCTTTGC
601 CGCCTGCG TGCTGATT TC AACCGCGAA GCCTTCGGC ATCTCTGCGT
```



```

651  CGGCACGGCG ATGGGTA1.. TGTTCGGCTT TGCTTTTGTG ATGATGACGC
701  TCAAAGGCTA CAGCAGCGCG GGGCATGTCT ATTCGGTCGG CACTTATCTG
751  TGGATGTTG CCATGAGTTT GGACGACGTG CCGCGATTGG TCGAACAATA
801  TTCCAATTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA
851  TCAAAGCCGG AACTTGA

```

This corresponds to the amino acid sequence <SEQ ID 34; ORF 006-1>:

```

m006-1.pep
  1  MWKMLKHIAQ THRKRLIGTF SLVGLENLLM LVYFVFGGRA INAVIAGEVW
 51  QALLYALVVL LMWLVGAVRR IADTRTFTRI YTEIAVPVVL EQRQRQVPHS
101  AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSavg
151  ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGD RR QLYRHYGLLA
201  RLRVLISNRE AFGYLCVGTA MGILFGFAFV MMTLKGYS SA GHVYSVGTYL
251  WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*

```

m006-1/g006-1 95.5% identity in 288 aa overlap

m006-1.pep	10	20	30	40	50	60
	MWKMLKHIAQTHRKRLIGTFSLVGLENLLMLVYFVFGGRAINAVIAGEVWQALLYALVVL					
g006-1	10	20	30	40	50	60
	MWKMLKHIAKTHRKRLIGTFSPVGLENLLMLGYFVFGGWAINAVIAGRvwQALLYALVVF					
m006-1.pep	70	80	90	100	110	120
	LMWLVGAVRR IADTRTFTRI YTEIAVPVVL EQRQRQVPHSAVTARVALSREFVSFFEEHL					
g006-1	70	80	90	100	110	120
	LMWLVGAAARR IADTRTFTRI YTEIAVPVVL EQRQRQVPHSAVTARVALSREFVSFFEEHL					
m006-1.pep	130	140	150	160	170	180
	PIAATSVVSI FGACIMLLVLEFWVGVSavgILALFLWLLPRFAAISENLYFRLNNSLERD					
g006-1	130	140	150	160	170	180
	PIAATSVVSI FGACIMLLVLEFWVGVSavgILALFLWLLPRFAAISENLYFRLNNSLERD					
m006-1.pep	190	200	210	220	230	240
	NHFIRKGD RR QLYRHYGLLARLRVLISNREAFGYLCVGTA MGILFGFAFVMMTLKGYS SA					
g006-1	190	200	210	220	230	240
	NHFIRKGD RR QLYRHYGLVSLRVLISNREAFGYLCVGAAMGILFGFAFVMMTLKGYS SA					
m006-1.pep	250	260	270	280	289	
	GHVYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX					
g006-1	250	260	270	280		
	GHIYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 35>:

```

a006-1.seq (partial)
  1  ..AGCCAAAACC ACCGCAAGCG ATTGATTGGC ACATTTTTC TGGTCCGACT
 51  GGAAAACCTT TTGATGCTGG TGTATCCGGT GTTTGGCGGC TGGCGGATTA
101  ATGCCGTGAT TCGGGGCGAG CCGTGGCAGG CGTTGCTGTA CGCTTTGGTT
151  GTGCTTTTGA TGTGGCTGGT CGGTGCGGCG CGGCGGATTG CCGATACCGG
201  CACGTTTACG CGGATTATA CCGAAATCGC CGTGCCGGTT GTGTTGGAAC
251  AGCGGCAGCG GCAAGTCCCG CATTGCGCGG TAACTGCGCG GGTGCGCTG
301  TCGCGTGAGT TTGTCAGCTT TTTGAAGAA CACCTGCCGA TTGCCGCGAC
351  ATCCGTGCTA TCCATATTCG GCGCGTGCAT CATGCTGCTG GTGCTGGAAT
401  TTTGGGTCGG CGTGTGCGCG GTGGGCATAC TTGCGTTGTT TTTATGGCTT
451  TTGCCACGTT TTGCCGCCAT CAGCGAAAAC CTGTATTTCC GCCTGAAGAA
501  CAGCTTGGAA CGCGACAACC ACTTTATCCG AAAAGGCGAC GAGCGGCAGC
551  TGGACCGCCA TTACGGACTG CTTGCGCGCC TGCGTGTGCT GATTTCCAAC
601  CGCGAAGCCT TCGGCTATCT CTGCGTCCGC ACGGCGATGG GTATTTTGTT
651  CGGCTTTGCT TTTGTGATGA TGACGCTCAA AGGCTACAGC AGCGCGGGGC

```

```

701 ATGTCTATTC GGTCCGCACT TATCTGTGGA TGTTTGCCAT AAGTTTGGAC
751 GACGTGCCGC GATTGGTCGA ACAATATTCC AATTGAAAG ACATCGGACA
801 ACGGATAGAG TGGTCGAAAC GGAACATCAA AGCCGGAAC TGA

```

This corresponds to the amino acid sequence <SEQ ID 36; ORF 006-1.a>:

```

a006-1.pep (partial)
1  ..SQNHRKRLIG TFFLVGLENL LMLVYPVFGG WAINAVIAGQ AWQALLYALV
51  VLLMWLVGAA RRIADTRTFT RIYTEIAVPV VLEQRQROVP HSAVTARVAL
101 SREFVSFFEE HLPAAATSVV SIFGACIMLL VLEFWVGVS A VGILALFLWL
151 LPRFAAISEN LYFRLKNSLE RDNHFIRKGD ERQLDRHYGL LARLRVLISN
201 REAFGYLCVG TAMGILFGFA FVMMTLKGYS SAGHVYSVGT YLWMFAISLD
251 DVPRLVEQYS NLKDIGQRIE WSKRNIKAGT *

```

a006-1/m006-1 95.7% identity in 280 aa overlap

	10	20	30	40	50
a006-1.pep	SQNHRKRLIGTFFLVGLENL LMLVYPVFGG WAINAVIAGQAWQALLYALVVL				
m006-1	MWKMLKHIAQTHRKRLIGTFFSLVGLENL LMLVYPVFGG RAINAVIAGEVWQALLYALVVL				
	10	20	30	40	50
	60	70	80	90	100
a006-1.pep	LMWLVGAAARRIADTRTFTRIYTEIAVPV VLEQRQROVP HSAVTARVALSREFVSFFEEHL				
m006-1	LMWLVGAVRRIADTRTFTRIYTEIAVPV VLEQRQROVP HSAVTARVALSREFVSFFEEHL				
	70	80	90	100	110
	120	130	140	150	160
a006-1.pep	PIAATSVVSIFGACIMLLVLEFWVGVS AVGILALFLWLLPRFAAISENLYFRLKNSLERD				
m006-1	PIAATSVVSIFGACIMLLVLEFWVGVS AVGILALFLWLLPRFAAISENLYFRLNNSLERD				
	130	140	150	160	170
	180	190	200	210	220
a006-1.pep	NHFIRKGDERRQLDRHYGLLARLRVLISN REAFGYLCVGTAMGILFGFAFVMMTLKGYS				
m006-1	NHFIRKGDERRQLDRHYGLLARLRVLISN REAFGYLCVGTAMGILFGFAFVMMTLKGYS				
	190	200	210	220	230
	240	250	260	270	280
a006-1.pep	GHVYSVGTLYLWMFAISLDDVPRLVEQYS NLKDIGQRIEWSKRNIKAGTX				
m006-1	GHVYSVGTLYLWMFAMSLDDVPRLVEQYS NLKDIGQRIEWSERNIKAGTX				
	250	260	270	280	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 37>:

```

g007.seq
1  atgaACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGcgC
51  CGCcGCTTCT GCCGccgaca acAGCatcat gaCaAAAGGG CAAAAAGTGT
101 ACGAATCcAa ctGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
151 ACTGCGtTTC CTccgctTTT CCggtcgGac tgtattatga acaAACCGCa
201 cgTCCTgctg cacagcatgg tcaaaggcAt cgacgggaca ttcaaagtgg
251 agcggcaaaa cctacgacgg atttatgCcc gcaaccgcca tcagcgATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 38; ORF 007.ng>:

```

g007.pep
1  MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKYESNCIA CHG*KKEGRG
51  TAFPPLFRSD CIMNKPHVLL HSMVKGIDGT FKVERQNLRR IYARNRHQRC
101 GHCRRRHLYH ERL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 39>:

**m007.seq**

```

1 ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251 AACGGCAAAA CCTACAACGG ATTCATGCCC GCAACCGCCA TCAGCGATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 40; ORF 007>:

**m007.pep**

```

1 MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARNRHQRC
101 GHCRRRHLYH ERL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 41>:

**a007.seq**

```

1 ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251 AACGGCAAAA CCTACAACGG ATTCATGCCC GCCACTGCCA TCAGCGATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 42; ORF 007.a>:

**a007.pep**

```

1 MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRED FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARHCHQRC
101 GHCRRRHLYH ERL*

```

**m007/a007** 97.3% identity over a 113 aa overlap

	10	20	30	40	50	60
m007.pep	MNTTRLPTALVLGCFCAAAS	ADNSIMTKGQKVYESNCV	ACHGKKGEGRG	TMFPPLYRSD		
a007	MNTTRLPTALVLGCLCAAAS	ADNSIMTKGQKVYESNCV	ACHGKKGEGRG	TMFPPLYRSD		
	10	20	30	40	50	60
	70	80	90	100	110	
m007.pep	FIMKKPQVLLHSMVKGINGT	IKVXRQNLQRIHARNRHQRC	GHCRRRHLYHERLX			
a007	FIMKKPQVLLHSMVKGINGT	IKVXRQNLQRIHARHCHQRC	GHCRRRHLYHERLX			
	70	80	90	100	110	

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 007 shows 86.7% identity over a 113 aa overlap with a predicted ORF (ORF 007.ng) from *N. gonorrhoeae*:

**m007/g007**

	10	20	30	40	50	60
m007.pep	MNTTRLPTALVLGCFCAAAS	ADNSIMTKGQKVYESNCV	ACHGKKGEGRG	TMFPPLYRSD		
g007	MNTTRLPTAFILCCLCAAAS	ADNSIMTKGQKVYESNCI	ACHGKKGEGRG	TAFFPLFRSD		
	10	20	30	40	50	60
	70	80	90	100	110	
m007.pep	FIMKKPQVLLHSMVKGINGT	IKVXRQNLQRIHARNRHQRC	GHCRRRHLYHERLX			
	70	80	90	100	110	

g007                    CIMNKPVLLHSMVKGIDGTFKVERQNLRRYARNRHQRCGHCRRRHLYHERL  
                               70                80                90                100                110

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 43>:

g007-1.seq (partial)  
 1 ATGAACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGCGC  
 51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT  
 101 ACGAATCCAA CTGCATCGCC TGCCACGGCA AGAAAGGGGA AGGCGCGGC  
 151 ACTGCGTTTC CTCCGCTTTT CCGGTCGGAC TATATTATGA ACAAACCGCA  
 201 CGTCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA  
 251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG  
 301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG  
 351 CGGAAGCGTT ACCGAAAAAG ACGTAAACA GGCAAAAGGC AAAAAAAC.

This corresponds to the amino acid sequence <SEQ ID 44; ORF 007-1.ng>:

g007-1.pep (partial)  
 1 MNTTRLPTAF ILCCLCAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG  
 51 TAFPPLFRSD YIMNKPVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA  
 101 DIAAVATYIM NAFDNGGGSV TEKDVQAKG KKN...

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 45>:

m007-1.seq  
 1 ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC  
 51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT  
 101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA  
 151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAAACCGCA  
 201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA  
 251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG  
 301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG  
 351 CGGAAGCGTT ACCGAAAAAG ACGTAAACA GGCAAAAGC AAAAAAACT  
 401 AA

This corresponds to the amino acid sequence <SEQ ID 46; ORF 007-1>

m007-1.pep  
 1 MNTTRLPTAL VLGCFCAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG  
 51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA  
 101 DIAAVATYIM NAFDNGGGSV TEKDVQAKS KKN\*

m007-1 / g007-1 91.7% identity in 133 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALV	LGCFCAASA	ADNSIMTKG	QKVYESNCV	ACHGKKGEG	RGTMFPLYRSD
g007-1	MNTTRLPTAF	ILCCLCAAS	AADNSIMTKG	QKVYESNCI	ACHGKKGEG	RGTAFPLFRSD
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPQVLL	HSMVKGING	TIKVNGKTY	NGFMPATAI	SDADIAAV	ATYIMNAFD
g007-1	YIMNKPVLL	HSMVKGING	TIKVNGKTY	NGFMPATAI	SDADIAAV	ATYIMNAFD
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKS	SKKNX				
g007-1	TEKDVQAK	GKKKN				
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 47>:

a007-1.seq (partial)  
 1 ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC  
 51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT  
 101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCA AGGCCGCGGA

```

151 ACCATGTTTC CGCCGCTCA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CCACTGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAAG ACGTAAACA GGCAAAAAAC AAAAAA..

```

This corresponds to the amino acid sequence <SEQ ID 48; ORF 007-1.a>:

**a007-1.pep (partial)**

```

1 MNTTRLPTAL VLGLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKN KK..

```

m007-1/a007-1 98.5% identity in 132 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGLCAAAS	ADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
a007-1	MNTTRLPTALVLGLCAAAS	ADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPVLLHSMVKGINGT	IKVNGKTYNGFMPATAISDA	DIAAVATYIMNAFDNGGGSV			
a007-1	FIMKKPVLLHSMVKGINGT	IKVNGKTYNGFMPATAISDA	DIAAVATYIMNAFDNGGGSV			
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKSKKNX					
a007-1	TEKDVQAKNKK					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 49>:

**g008.seq**

```

1 ATGAACAACA GACATTTTGC CGTCATcgCC TTGGGCAGCA ACCTTGACAA
51 CCCCACACAA CAAATacgcg gcgcattaga cgcgctctcg tcccatcctg
101 acatccggct tgaacaggtt tcctcactgt aTatgaccgc acctgtcggT
151 tacgAcaaTC agcccgATTT CATCaatgcc gTCTgcaccg TTTCACCAC
201 CtTGGACGGC ATTGcccTGC TTGCCgaACT CAAccgTATC GAAGCCGATT
251 TCGGACGCGA aCGCAGTTTC CGCAATGCAC CGCGCACATT GGATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGCC TTACCCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAAATACG GAAAGTTGT CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGACA GGTAA

```

This corresponds to the amino acid sequence <SEQ ID 50; ORF 008.ng>:

**g008.pep**

```

1 MNNRHFAVIA LGSNLDNPAQ QIRGALDALS SHPDIRLEQV SSLYMTAPVG
51 YDNQPDFINA VCTVSTTLDG IALLAELNRI EADFGRRERSF RNAPRTLDD
101 IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKYGKVVELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 51>:

**m008.seq**

```

1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
51 CCCTGCTCAA CAGGTACGCG CCGCATTGGA CACGCTGTCTG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT TGTCaATGCC GTCTGCACCG TTTCACCAC
201 TCTGGACGGC ATTGCCyTGC TTGCCGAAct CAACCGTATC GAGGCTGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GkATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACACsCGAC TcACcTtGCC

```

```

351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATCCGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTGTTTTA GGAAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGyTGG GCAATCAAGG TATCCGTCTT TTACCGGACA GGTAAAT

```

This corresponds to the amino acid sequence <SEQ ID 52; ORF 008>:

```

m008.pep
1  MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
51  YDNQPDFVNA VCTVSTLDG IALLAELNRI EADFGRRERSF RNAPRTLXLD
101 IIDFDGISSD DTRLTLPHPR AHERSFVIRP LAEILPDFVL GKHKVAELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis*<SEQ ID 53>:

```

a008.seq
1  ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTGAAAA
51  CCCTGCCCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT CGTCAATGCC GTCTGCACCG TTTCCACCAC
201 CTTGGACGGC ATTGCCCTGC TTGCCGAACT CAACCGTATC GAAGCCGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GGATTTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGGAC TCACCCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGATA AGTAA

```

This corresponds to the amino acid sequence <SEQ ID 54; ORF 008.a>:

```

a008.pep
1  MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG
51  YDNQPDFVNA VCTVSTLDG IALLAELNRI EADFGRRERSF RNAPRTLXLD
101 IIDFDGISSD DTRLTLPHPR AHERSFVIRP LAEILPDFIL GKHKVAELS
151 KRLGNQGIRL LPDK*

```

m008/a008 97.6% identity over a 164 aa overlap

	10	20	30	40	50	60
m008.pep	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
a008	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
	10	20	30	40	50	60
m008.pep	VCTVSTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
a008	VCTVSTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
	70	80	90	100	110	120
m008.pep	VCTVSTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
a008	VCTVSTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
	70	80	90	100	110	120
m008.pep	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
a008	AHERSFVIRPLAEILPDFILGKHGKVAELSKRLGNQGIRLLPDKX					
	130	140	150	160		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 008 shows 92.7% identity over a 164 aa overlap with a predicted ORF (ORF008.ng) from *N. gonorrhoeae*:

m008/g008

	10	20	30	40	50	60
m008.pep	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
g008	MNNRHFAVIALGSNLDNPAQQIRGALDALSSHPDIRLEQVSSLYMTAPVGYDNQPDFINA					

183

	10	20	30	40	50	60
m008.pep	70	80	90	100	110	120
	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
g008	70	80	90	100	110	120
	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLDLDIIDFDGISSDDPRLTLPHPR					
m008.pep	130	140	150	160		
	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
g008	130	140	150	160		
	AHERSFVIRPLAEILPDFILGKYGVVVELSKRLGNQGIRLLPDRX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 55>:

**g009.seq**

```

1  ATGCCCCGCG CTGCCGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GCGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAA TCAGCACACG CAGGCGCGAA ACCAGTCCGT AATGGCGGTA
151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTcg cGTCCAAGC
201 TGTTGTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AaaaGCCATA A

```

This corresponds to the amino acid sequence <SEQ ID 56; ORF 009.ng>:

**g009.pep**

```

1  MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFVGNQHT QARNQSVMAV
51  QLPLVAFSDK VVVAFAQVVQ AEIQVFADGG KTWQKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 57>:

**m009.seq**

```

1  ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTA
151 CAGCTGCCGC CGGTCGCCTT TTCGGATAAA GTGGTTGTcg CGTCCAAGC
201 TGTTGTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A

```

This corresponds to the amino acid sequence <SEQ ID 58; ORF 009>:

**m009.pep**

```

1  MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFVGNQHT QARKQSVMAV
51  QLPPVAFSDK VVVAFAQVVQ AEIQVFADGG KTWQKP*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 009 shows 97.7% identity over a 86 aa overlap with a predicted ORF (ORF 009.ng) from *N. gonorrhoeae*:

**m009/g009**

	10	20	30	40	50	60
m009.pep	MPRAAVAFERHHHKSKEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK					
g009	10	20	30	40	50	60
	MPRAAVAFERHHHKSKEQNTHRRADAEIAEGFAVGNQHTQARNQSVMAVQLPLVAFSDK					
m009.pep	70	80				
	VVVAFAQVVQAEIQVFADGGKWTWQKPx					
g009	70	80				
	VVVAFAQVVQAEIQVFADGGKWTWQKPx					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 59>:

```
a009.seq
1  ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTC
151 CAGCTGCCGC TCGTCGCCTT TTCGATAAA GTGGTTGTCG CGTTCCAAGC
201 TGTTCCTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 60; ORF 009.a>:

```
a009.pep
1  MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFVGNQHT QARKQSVMAV
51  QLPLVAFSDK VVVAFAQVLQ AEIQVFADGG KTWQKP*
```

m009/a009 97.7% identity over a 86 aa overlap

```

          10      20      30      40      50      60
m009.pep  MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGT  ENQHTQARKQSVMAVQLPPLVAFSDK
          |||||
a009      MPRAAVAFERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPLVAFSDK
          10      20      30      40      50      60

          70      80
m009.pep  VVVAFAQVAVQAEIQVFADGGKTWQKPx
          |||||
a009      VVVAFAQVLQAEIQVFADGGKTWQKPx
          70      80
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 61>:

```
g010.seq
1  ATGGGTTTTC CTGTTTCGCA GTTTGATGCC GTGATTGTCG GCGGTGGCGG
51  TGCAGGTTTA CGTGACGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTGTGC TAAAGTGTTC CCGACCGGCT CGCATACCGT AGCGGCGCAG
151 GCGGATATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTGC TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGTTTGG GCATTTGCGC CCGTGCGGGC ATTCGGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGGAG GCGGTTATTC TGTTGAacgc
801 cgacggcgaa cgcTTTATGG AAcgctatgc GCcgACCGta aAagaCTTGG
851 CTTCTCGCga cgtGGTTTCA CgcgcGatgG CGatggaAat ctatgaaggt
901 cgcggctgTG GtaaaAAcaA agaCCacgtC TTA CTGAAAA TCGACcAtAt
951 cggTGCAGAA AAAATTATGG AAAAATGCC GGGCATCCGC GAGATTTCa
1001 TTCagtttgc cGGTATCGAT CCGATTAAAG ACCCGATTcc ggttgTGCCG
1051 ACTACCCACT ATATGATGGG CGGCATTcGg aCCAATTATC ACGGTGAAGT
1101 TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT
1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga
```

This corresponds to the amino acid sequence <SEQ ID 62; ORF 010.ng>:

```
g010.pep
```



```

1  MGFVPRKFDA VIVGGGGAOL RAALQLSKSG LNCVLSKV F PTRSHTVAAQ
51  GGISASLGNV QEDRWDRWHY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPEFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNV RANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFRAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCCKNKDHF LLKIDHIGAE KIMEKLP GIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEVVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTPR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 63>:

```

m010.seq (PARTIAL)
1  ..nTCCAATTAT CCAAATCCGG TCTGAATTGT GCGGTTTTGT CTAAAGTGTT
51  CCCGACCCGT TCGCATACCG TAGCGGCGCA GGGCGGTATT TCCGCCTCTn
101 TGGGTAATGT GCAGGAAGAC CGTTGGGACT GGCACATGTA CGATACCGTG
151 AAAGGTTCCG ACTGGTTGGG CGACCAAGAT GCGATTGAGT TTATGTGCCG
201 CGCCGCGCCT GAAGCCGTAA TTGAGTTGGA ACACATGGGT ATGCCTTTTG
251 ACCGTGTGGA AAGCGGTAAA ATTTATCAGC GTCCTTTCCG CGGCCATACT
301 GCCGAACACG GTAAACGCGC GGTAGAACGC GYCTGTGCGG TTGCCGACCG
351 TACAGGTCAT GCGATGCTGC ATACTTTGTA CCAACAAAAC GTCCGTGCCA
401 ATACGCAATT CTTTGTGGAA TGGACGGCAC AAGATTTGAT TCGTGATGAA
451 AACGGCGATG TCGTCGGCGT AACCGCCATG GAAATGGAAA CCGGCGAAgt
501 TTATATTTTC CACGCTAAAG CTGTGATGTT TGCTACCGGC GCGGCGGTC
551 GTATTTATGC GTCTTCTACC AATGCCTATA TGAATACCGG CGATGGTTTG
601 GGTATTTGTG CGCGTGCAGG TATCCCCTTG GAAGACATGG AATTCTGGCA
651 ATCCAGCCG ACCGGCGTGG CGGGTGCGGG CGTGTGATT ACCGAA...

```

This corresponds to the amino acid sequence <SEQ ID 64; ORF 010>:

```

m010.pep (PARTIAL)
1  ..XQLSKSGLNC AVLSKVFPT R SHTVAAQGGI SASXGNVQED RWDWHMYDTV
51  KGSDWLGDQD AIEFMCRAP EAVIELEHMG MPFDRVESGK IYQRPFGGHT
101 AEHGKRAVER XCAVADR TGH AMLHTLYQQN VRANTQFFVE WTAQDLIRDE
151 NGDVVGVTAM EMETGEVYIF HAKAVMFATG GGGRIYASST NAYMNTGDGL
201 GICARAGIPL EDMFEWQFQP TGVAGAGVLI TE...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 65>:

```

a010.seq
1  ATGGGCTTTC CTGTTGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCG GCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCCGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
451 CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
751 TGTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGATATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTCTCTCGCA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTAAGTAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAAAGTGGC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTTCG CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCGG
1051 ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
1101 TGTCGTTTCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTGGGTG
1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAGAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA

```

```

1301 ATGCCGGCGA ACTGACCCGC CAACGTATCG AGCGTTTGA CAATCAAAC
1351 GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401 ATTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTAATTGAA GTGGCGAAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCACACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCTTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 66; ORF 010.a>:

```

a010.pep
1  MGFVPRKFDA VIVGGGGAGL RAXLQLSKSG LNCVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDWHMY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMFPDRVE SGKIYQRPFG GHAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNV RANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCCKNDHV LLKIDHIGAE KIMEKLP GIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVPV QGDEYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLDDL VVF GKAAGDSMIK FIKEQSDWK P LPANAGELTR QRIERLNDQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVM AIA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRG AHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*

```

m010/a010 98.7% identity over a 231 aa overlap

```

                                10      20      30
m010.pep                      XQLSKSGLNCVLSKVF PTRSHTVAAQGGISASXGNV
                                |||
a010      MGFVPRKFDAVIVGGGGAGLRAXLQLSKSG LNCVLSKVF PTRSHTVAAQGGISASLGNV
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m010.pep      QEDRWDWHMYDTVKGSDWLG DQDAIEFMCR AAPEAVIELE HMGMFPDRVESGKIYQRPFG
a010      QEDRWDWHMYDTVKGSDWLG DQDAIEFMCR AAPEAVIELE HMGMFPDRVESGKIYQRPFG
                                70      80      90      100     110     120

                                100     110     120     130     140     150
m010.pep      GHAEHGKRAVERXCAVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVG V
a010      GHAEHGKRAVERACAVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVG V
                                130     140     150     160     170     180

                                160     170     180     190     200     210
m010.pep      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIP LEDMEFWQ
a010      TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIP LEDMEFWQ
                                190     200     210     220     230     240

                                220     230
m010.pep      FQPTGVAGAGV LITE
a010      FHPTGVAGAGV LITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVS RAMAMEIYEG
                                250     260     270     280     290     300

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 010 shows 98.7% identity over a 231 aa overlap with a predicted ORF (ORF 010.ng) from *N. gonorrhoeae*:

m010.pep/g010.pep

m010.pep				10	20	30		
				XQLSKSGLNCAVL	SKVFPTRSH	TVAAQGGISAS	XGNV	
g010	MGFPVRKFDA	VIVGGGAGL	RAALQLSKSGLNCAVL	SKVFPTRSH	TVAAQGGISAS	LGNV		
	10	20	30	40	50	60		
m010.pep		40	50	60	70	80	90	
		QEDRWDWHMYD	TVKGS	DWLGDQDAIE	FMCR	AAPEAVIE	LEHMGMPFDR	VESGKIYQRPFG
g010		QEDRWDWHMYD	TVKGS	DWLGDQDAIE	FMCR	AAPEAVIE	LEHMGMPFDR	VESGKIYQRPFG
		70	80	90	100	110	120	
m010.pep		100	110	120	130	140	150	
		GHTAEHGKRA	VERXCAVAD	RTGHAMLHT	LYQQNV	RANTQFF	VEWTAQDL	IREDNGDVVGV
g010		GHTAEHGKRA	VERACAVAD	RTGHAMLHT	LYQQNV	RANTQFF	VEWTAQDL	IREDNGDVVGV
		130	140	150	160	170	180	
m010.pep		160	170	180	190	200	210	
		TAMEMETGEV	YIFHAKAV	MFATGGG	GRIYASST	NAYMNTG	DGLGICAR	AGIPL
g010		TAMEMETGEV	YIFHAKAV	MFATGGG	GRIYASST	NAYMNTG	DGLGICAR	AGIPL
		190	200	210	220	230	240	
m010.pep		220	230					
		FQPTGVAGAG	VGLITE					
g010		FHPTGVAGAG	VGLITEGVR	GEGGILLN	ADGERFM	ERYAPT	VKDLASR	
		250	260	270	280	290	300	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 67>:

g010-1.seq..

```

1  ATGGGTTTTC CTGTTCCGAA GTTTGATGCC GTGATTGTCTG GCGGTGGCGG
51  TGCAGGTTTA CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTC CCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTGATGAT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTCTG TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCGGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGTATTC TGTTGAACGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTGGTTTCA CCGCGCATGG CGATGGAAT CTATGAAGGT
901 CGCGGCTGTG GTAAAAACAA AGACCACGTC TTAAGTAAAA TCGACCATAT
951 CCGTGCAGAA AAAATTATGG AAAAAGTACC GGCATCCGC GAGATTTCCT
1001 TTCAGTTTGC CGGTATCGAT CCGATTAAAG ACCCGATTCC GGTGTGCCC
1051 ACTACCCACT ATATGATGGG CGGCATCCG ACCAATTATC ACGGTGAAGT
1101 TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAATA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT

```

1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTc cgcccaacccc cccggtga

This corresponds to the amino acid sequence <SEQ ID 68; ORF 010-1.ng>:

g010-1.pep

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDWHMY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMFPDRVE SGKIQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNV RANTQF FVEWTAQDLI RDENG DVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLASRDVVS RAMAMEIYEG
301 RGC GKNKDHV LLKIDHIGAE KIMEKLP GIR EISIQFAGID PIKQPIPVVP
351 TTHYMMGGIP TNYHGEVVPV QGDEYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTPR*

```

g010-1 / P10444

sp|P10444|DHSA\_ECOLI SUCCINATE DEHYDROGENASE FLAVOPROTEIN SUBUNIT  
 gnl|PID|d1015210 (D90711) Succinate dehydrogenase, flavoprotein [Escherichia coli] gi|1786942  
 (AE000175) succinate dehydrogenase flavoprotein subunit [Escherichia coli] Length = 588  
 Score = 1073 (495.6 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169  
 Identities = 191/303 (63%), Positives = 238/303 (78%)

Query: 1 MGFPVRKFDAVIVXXXXXXXXXXXXXSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV 60  
 M PVR+FDVAV++ S+SG CA+LSKVFPTRNTV+AQGGI+ +LGN  
 Sbjct: 1 MKLPVREFDAVVIGAGGAGMRAALQISQSGQTCALLSKVFPTRSHTVSAQGGITVALGNT 60

Query: 61 QEDRWDWHMYDTVKGSDWLGQDAIEFMCRAAPEAVIELEHMGMFPDRVESGKIYQRPFG 120  
 ED W+WHMYDTVKGS++GDQDAIE+MC+ PEA++ELEHMG+PF R++ G+IYQRPFG  
 Sbjct: 61 HEDNWEWHMYDTVKGS DYIGDQDAIEYMCKTGPEAILEHMG LPPFSRLDDGRIYQRPFG 120

Query: 121 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVGV 180  
 G + G E A ADRTGHA+LHTLYQQN++ +T F EW A DL+++++G VVG  
 Sbjct: 121 GQSKNFGGEQAARTAAADRTGHALLHTLYQQNLKXHTTIFSEWYALDLVKNQDGA VVGC 180

Query: 181 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTG DGLGICARAGIPLEDMEFWQ 240  
 TA+ +ETGEV F A+A + ATGG GRIY S+TNA++NTGDG+G+ RAG+P++DME WQ  
 Sbjct: 181 TALCIETGEVYFKARATVLTATGGAGRIYQSTTNAHINTGDGVGM AIRAGVPVQDMEMWQ 240

Query: 241 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPT VKDLASRDVVS RAMAMEIYEG 300  
 FHPTG+AGAGVL+TEG RGE G LLN GERFMERYAP KDLA RDVV+R++ +EI EG  
 Sbjct: 241 FHPTGIAGAGVLVTEGCRGEGGYLLNKHGERFMERYAPNAKDLAGRDVVARSIMIEIREG 300

Query: 301 RGC 303  
 RGC  
 Sbjct: 301 RGC 303

Score = 249 (115.0 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169  
 Identities = 53/102 (51%), Positives = 62/102 (60%)

Query: 309 HVLLKIDHIGAEKIMEKLP GIREISIQFAGXXXXXXXXXXXXTTHYMMGGIPTNYHGEV 368  
 H LK+DH+G E + +LPGI E+S FA T HYMMGGIPT G+ +  
 Sbjct: 310 HAKLKLDHLGKEVLESRLPGILELSRTFAHVD PVKEPIPVIP TCHYMMGGIPTKVGTGQAL 369

Query: 369 VPQGDEYEV PVKGLYAAGECACASVHGANRLGTNSLLDLVVF 410  
 +V V GL+A GE AC SVHGANRLG NSLLDLVVF  
 Sbjct: 370 TVNEKGEDVVVPGLFAVGEIACVSVHGANRLGNSLLDLVVF 411

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 69>:

m010-1.seq..

```

1  ATGGGTTTTC CTGTTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCAGCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTGTGC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GCGGTATTT CCGCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCGGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCGGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC

```

```

451 CAACAAAACG TCCGTGCGAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGTGCG TATTATATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTGTGTC GCGTGCAGGT ATCCCGTTGG
701 AAGACAT3GA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGCGGT ACGGCGCGAG GCGGATATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAAATGCC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTGTGCGCG
1051 ACTACCCACT ATATGATGGG CGGCATTCCG ACCAAATACC ACGGCGAAGT
1101 TGTCGTTCGG CAAGGTGAAG ATTACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCTTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACCAACTCCC TGTGGACTT GGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAT TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCTTGCTA
1301 ATGCAGGTGA GTTGACCCGC CAACGTATCG AGCGTTTGA CAACCAAAACC
1351 GATGGTGAAA ACGTTGATGC ATTGCGTCGC GAACTGCAAC GCTCTGTACA
1401 ACTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGCGCTTC
1451 GAGAAGTCAT GCGGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTGATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGGCGACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATATCAATA CCTGTCCCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGATGA

```

This corresponds to the amino acid sequence <SEQ ID 70; ORF 010-1>:

m010-1.pep..

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVFE PTRSHTVAAQ
51  GGISASLGNV QEDRWDMHYD DTVKGSDDLW GDQDAIEFMCRA APEAVIELEH
101 HMGMPPDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLSARDVVS RAMAMEIYEG
301 RCGCKNKDHF LLKIDHIGAE KIMEKLPGR EISIQFAGID PIKDFIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGEDYEVVVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIRLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMIA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DINTLSYKPV HTKPLSVEYI KPAKRVY*

```

m010-1 / g010-1 99.5% identity in 410 aa overlap

	10	20	30	40	50	60
m010-1.pep	MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCVLSKVFPTRSHTVAAQGGISASLGNV					
g010-1	MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCVLSKVFPTRSHTVAAQGGISASLGNV					
	10	20	30	40	50	60
m010-1.pep	QEDRWDMHYDITVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPDRVESGKIYQRPFG					
g010-1	QEDRWDMHYDITVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPDRVESGKIYQRPFG					
	70	80	90	100	110	120
m010-1.pep	GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDVVGV					
g010-1	GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDVVGV					
	130	140	150	160	170	180
m010-1.pep	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
g010-1	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
	190	200	210	220	230	240
m010-1.pep	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
g010-1	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					

190

	190	200	210	220	230	240
m010-1.pep	250	260	270	280	290	300
	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVS RAMAMEIYEG					
g010-1	250	260	270	280	290	300
	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVS RAMAMEIYEG					
m010-1.pep	310	320	330	340	350	360
	RGCGKNKDHVLLKIDHIGA EKIMEKLP GIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP					
g010-1	310	320	330	340	350	360
	RGCGKNKDHVLLKIDHIGA EKIMEKLP GIREISIQFAGIDPIKDPIPVVPTTHYMMGGIP					
m010-1.pep	370	380	390	400	410	420
	TNYHGEVVVPQGDEYEVVKGLYAAGECACASVHGANRLGTNSLLDLVVFKAAGDSMIK					
g010-1	370	380	390	400	410	
	TNYHGEVVVPQGDEYEVVKGLYAAGECACASVHGANRLGTNSLLDLVVFRTPRX					
m010-1.pep	430	440	450	460	470	480
	FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTDEILS					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 71>:

a010-1.seq..

```

1  ATGGGCTTTC CTGTTCCGCAA GTTGTATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCTCTCTT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
451 CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGCGCGCAG GCGGATATTC TGTGTAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCGGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAAAGTACC GGCATCCGC GAGATTTCCT
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051 ACTACCCACT ATATGATGGG CCGTATTCGG ACCAACTACC ATGGCGAAGT
1101 TGTCGTTTCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAACCTT TTGCCTGCTA
1301 ATGCCGCGCA ACTGACCCGC CAACGTATCG AGCGTTTGGA CAATCAAACCT
1351 GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401 ATGACGCGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GCGGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAG
1601 AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATCA GATGCAATA CTTGTCTCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 72; ORF 010-1.a>:

a010-1.pep..

```

1  MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDDHMY DTVKGSIDLW GDDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHTEHGKRA VERACAVADR TGHAMLHTLY

```

```
151 QQNVRRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCGKNKNDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPVPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEVVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKPLPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMAIA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*
```

m010-1 / a010-1 99.3% identity in 587 aa overlap

```
10 20 30 40 50 60
a010-1.pep MGFPVRKFDVAVIVGGGGAGLRAXLQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
|||||
a010-1 MGFPVRKFDVAVIVGGGGAGLRAXLQLSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV
10 20 30 40 50 60

70 80 90 100 110 120
a010-1.pep QEDRWDDHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
|||||
m010-1 QEDRWDDHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
70 80 90 100 110 120

130 140 150 160 170 180
a010-1.pep GHTEAHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDVVGV
|||||
m010-1 GHTEAHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDVVGV
130 140 150 160 170 180

190 200 210 220 230 240
a010-1.pep TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
|||||
m010-1 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
190 200 210 220 230 240

250 260 270 280 290 300
a010-1.pep FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVS RAMAMEIYEG
|||||
m010-1 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVS RAMAMEIYEG
250 260 270 280 290 300

310 320 330 340 350 360
a010-1.pep RGCGKNKNDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPVPVPTTHYMMGGIP
|||||
m010-1 RGCGKNKNDHVLLKIDHIGAEKIMEKLPGIREISIQFAGIDPIKDPVPVPTTHYMMGGIP
310 320 330 340 350 360

370 380 390 400 410 420
a010-1.pep TNYHGEVVVPQGDEYEVVKGLYAAGECACASVHGANRLGTNSLLDLVVFGKAAGDSMIK
|||||
m010-1 TNYHGEVVVPQGEDYEVVKGLYAAGECACASVHGANRLGTNSLLDLVVFGKAAGDSMIK
370 380 390 400 410 420

430 440 450 460 470 480
a010-1.pep FIKEQSDWKPLPANAGELTRQRIERLDNQT DGENVDALRRELQRSVQLHAGVFRTEILS
|||||
m010-1 FIKEQSDWKPLPANAGELTRQRIERLDNQT DGENVDALRRELQRSVQLHAGVFRTEILS
430 440 450 460 470 480

490 500 510 520 530 540
a010-1.pep KGVREVMAIAERVKRTEIKDKSKVWNTARIEALELDNLIEVAKATLVSAEARKESRGAHA
|||||
m010-1 KGVREVMAIAERVKRTEIKDKSKVWNTARIEALELDNLIEVAKATLVSAEARKESRGAHA
490 500 510 520 530 540

550 560 570 580
a010-1.pep SDDHPERDDENWMKHTLYHSDANTLSYKPVHTKPLSVEYIKPAKRVYX
|||||
```

m010-1                    SDDHPERDDENWMKHTLYHSDINTLSYKPVHTKPLSVEYIKPAKRVYX  
                               550                560                570                580

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 73>:

g011.seq  
 1 ATGAAGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC  
 51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA  
 101 GCCTGAAAAC CCGCCTTACC GAAGATATGA AAACCGCGAT GCGCGCCAAA  
 151 GATCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAATGCCG CCGTCAAAACA  
 201 GTTTGAAGTA GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA  
 251 TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACGCCGCGAA AATCTACACT  
 301 GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGACGT  
 351 GCTGCACCGC TACCTGCCGC AAATGCTCTC CGCCGGCGAA ATCCGCACCG  
 401 CCGTCGAAGC AGCCGTTGCC GAAACCGGCG CGGCAGGTAT GCGGGATATG  
 451 GGCAAAGTGA TGGTCGTATT GAAAAcccGC CTGCGCGGCA AAGccgATAT  
 501 GGGCGAAGTC AACAAAATCT TGAAAAccGt aCTGACCGCC tga

This corresponds to the amino acid sequence <SEQ ID 74; ORF 011.ng>:

g011.pep  
 1 MKTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKTRLT EDMKTAMRAK  
 51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDGAKIYT  
 101 EAGRQDLADK ENAEIDVLHR YLPQMLSAGE IRTAVEAAVA ETGAAGMADM  
 151 GKVMVVLKTR LAGKADMGEV NKILKTVLTA \*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 75>:

m011.seq (partial)  
 1 ATGAGGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC  
 51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA  
 101 GCCTGAAAAT CCGCCTTACC GAAGACATGA AAACCGCGAT GCGCGCCAAA  
 151 GACCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAACGCCG CCGTCAAAACA  
 201 GTTTGAAGTG GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA  
 251 TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACAGCGCGAA AATCTACACT  
 301 GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGAGGT  
 351 ACTGCACCGC TACCTTCCCC AAATGCTTTC CGCCGGCGAA ATCCGTACCG  
 401 AGGTCAAGC TGGCGTTGCC GAAACCGGCG CGGCAGGTAT GCGGGATATG  
 451 GGTAAGTCA TGGGGCTGCT GAAAACCCGC CTCGCAGGTA AAGCCGA...

This corresponds to the amino acid sequence <SEQ ID 76; ORF 011>:

m011.pep (partial)  
 1 MRTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKIRLT EDMKTAMRAK  
 51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDSAKIYT  
 101 EAGRQDLADK ENAEIEVLHR YLPQMLSAGE IRTEVEAAVA ETGAAGMADM  
 151 GKVMGLLKTR LAGKA.....

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 011 shows 95.8% identity over a 165 aa overlap with a predicted ORF (ORF 011.ng) from *N. gonorrhoeae*:

m011/g011

	10	20	30	40	50	60
m011.pep	MRTHRKTCSA	VCFAFQTASK	PAVSIRHPSE	DIMSLKIRLT	EDMKTAMRAK	DQVSLGTIRL
g011						
	10	20	30	40	50	60
m011.pep	INAAVKQFEV	DERTEADDAK	ITAILTKMVK	QRKDSAKIYT	TEAGRQDLADK	ENAEIEVLHR
g011						
	70	80	90	100	110	120
m011.pep	INAAVKQFEV	DERTEADDAK	ITAILTKMVK	QRKDGAKIYT	TEAGRQDLADK	ENAEIDVLHR
g011						



193

	70	80	90	100	110	120
	130	140	150	160		
m011.pep	YLPQMLSAGEIRTEVEAAVAETGAAGMADM	GKVMGLLKTRLAGKA				
g011	YLPQMLSAGEIRTA	VEAAVAETGAAGMADM	GKVMVVLKTRLAGKADM	GEVKNILKTVLTA		
	130	140	150	160	170	180
g011	X					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 77>:

**g012.seq**

```

1  ATGCTCGCCC GTCGCTATTT TTTCAATATC CAACCCGGGG CGGTTTTCAC
51  TGACAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGCCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACa
201 gGcggtGGAT ATTCGgcact tccgCcacca caccacccga accgatgacc
251 gcaaacggaG CGGAACAAT TTTATCCGCC acacacgcca tcatatagcc
301 gcCGCTTGCC GCGACCTTAT CGAcggcgac ggTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAATTCGCC GTCAAATCC AACACAAAAA GGCGGGATT TTTGCGTTTCG
551 GCAGATTCTC CCCCGCCCTC CTTCAAACGC TTTTCTCTCG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CTTTCTTTT CTGATGTTT GTCTCTTCCT
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 78; ORF 012.ng>:

**g012.pep**

```

1  MLARRYFFNI QPGAVFTDKL LEQLMRFLQF LPEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRFRHHHTR TDRKRSGNN FIRHRRHIA
101 AACRDLDIGD QQRNIAFAQT PKLRSRQTVT VNHAARTFQS EQNLIFRLGN
151 QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFRGRFLPAL LQTLFLCFGF
201 RLFLFLFFFF LMFCLFLA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 79>:

**m012.seq**

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 GCCGCTCGCn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
351 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
401 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
451 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
501 nnnnnnnnnn nnnnnnnnnC AACACAAAAA GGCGTGATT nTGCCTTTTCG
551 GCAGATTCTC CCCCACCTC CTTCAAACGT TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTGT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 80; ORF 012>:

**m012.pep**

```

1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRHHHTR TDNRKRSGN FIRHRRHIT
101 AARXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
151 XXXXXXXXXXX XXXXXXXXXXX XXXQHKKA*F XRFRGRFLPTL LQTFFLCFGF
201 RLFLFLFLFF LMLCLFPA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 81>:

```
a012.seq
1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTGCGAAT
101 TTCTGTTTGC CCTTTTCGGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTGCGCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GGCGGTGGAT ATTGCGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTGAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCCTAACC GTGAACCACG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACG TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GCGGGGATTT TTGCGTTTCG
551 GAAGATTCTC CCCACCCCTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
651 CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 82; ORF 012.a>:

```
a012.pep
1  MLARCHFLNI QLRVLAADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRYNTHR TDNRKRS GNN FIRHTRHHIT
101 TARRHLIDGD QORNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCLEFPA*
```

m012/a012 64.2% identity over a 218 aa overlap

m012.pep	10	20	30	40	50	60
	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
a012	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
	10	20	30	40	50	60
m012.pep	70	80	90	100	110	120
	NIMFFQQA VDI RYFRHHTHRTDNRKRS GSNFIRHTRHHITAA RXXXXXXXXXXXXXXXXXXXX					
a012	NIMFFQQA VDI RYFRYNTHRTDNRKRS GNNFIRHTRHHITARRHLIDGDGQORNIAFAQT					
	70	80	90	100	110	120
m012.pep	130	140	150	160	170	180
	XXXQHKKAXF					
a012	PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF					
	130	140	150	160	170	180
m012.pep	190	200	210	219		
	XRFGRLPTLLQTLFLCFGRFLFLFLFLFLMLCLFPAX					
a012	LRFGRLPTLLQTLFLCFGRFLFLFLFLFLFLMFCLEFPA					
	190	200	210			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 012 shows 58.7% identity over a 218 aa overlap with a predicted ORF (ORF 012.ng) from *N. gonorrhoeae*:

m012/g012

m012.pep	10	20	30	40	50	60
	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
g012	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					

```
m012-i.seq
1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTF CCTCCAGTTC CTGTGCGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTGATCA CATCCACATC AATATCATGT TTTTTCACAA
201 GCCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
251 GCAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGGCA TCATATAACC
301 GCCGCTCGCC GCCACCTTAT CGACGGCGAC GGTGAGCGGA ATATTGCGTT
351 CGCGCAAACG CyTAAGCTGC GAAGCGCCCA AACCGTAACC GTGAACACAG
401 CGCGCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTACG GCGTGGCAAT
451 CAAAAGCACC GCCCTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GCAGATTTCT CCCCACCCTC CTTCAAACGC TTTTCTCTCG CTTTGGCTTC
601 GCCTTTTTC TTTTCTTTT CCTCTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA
```

```
m012-1.pep
1  MLARCHFLNI  QLRVLADK  LEQLMRFLQ  LSEFLFALR  IFTHKSNRAL
51  KFARRHHIHI  NIMFFQAVD  TRYFRHHTH  TDNKRSGSN  FIRHTRHIT
101 AARRHLIDGD  GQRNIAFAQ  TKLRSRCTV  VNHAARTFS  EQNLIFRLGN
151 QKHRRLMTQ  GFYGVCIQA  VKIQHKKAG  LREGRLPTL  LQTLFLCFGF
201 RLFLFLFF  LMFCLFPA*
```

	10	20	30	40	50	60
m012-1.pep	MLARCHFLNIQLRAVLADKLLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKKFARRHHIHI					
g012	:  :      ::					
	10	20	30	40	50	60
	70	80	90	100	110	120
m012-1.pep	NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARRHLIDGDOGQRNIAFAQT					
g012	:            :      :      :					
	70	80	90	100	110	120
	130	140	150	160	170	180
m012-1.pep	XKLRSRQTVTVNHAAARTFQSEQNLIFRLGNQKHRRNLMTQGFGVCIQIAVKIQHKKAGF					
g012						
	PKLRSRQTVTVNHAAARTFQSEQNLIFRLGNQKHRRNLMTQGFGVCIQIAVKIQHKKAGF					

196

	130	140	150	160	170	180
	190	200	210	219		
m012-1.pep	LRFGRLPTLLQTLFLCFGRFLFLFLFLFLMFLCLFPAX					
g012	LRFGRLFPALLQTLFLCFGRFLFLFLFLFLMFLCLFLAX					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 85>:

a012-1.seq

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GCGGTGGAT ATTGGTACT TCCGTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACACG
401 CCGCCCGGAC TTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GCGGGGATTT TTGCGTTTCG
551 GAAGATTCTC CCCACCCCTC CTTCAAACGC . TTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 86; ORF 012-1.a>:

a012-1.pep

```

1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
101 TARRHLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKRRRLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCLFPA*

```

a012-1/m012-1 97.2% identity in 218 aa overlap

	10	20	30	40	50	60
a012-1.pep	MLARCHFLNIQLRAVLADKLL	LEQLMRFLQFLSEFLFALFR	IFTHKSNRALKFARRHHIHI			
m012-1	MLARCHFLNIQLRAVLADKLL	LEQLMRFLQFLSEFLFALFR	IFTHKSNRALKFARRHHIHI			
	10	20	30	40	50	60
	70	80	90	100	110	120
a012-1.pep	NIMFFQQA VD IRYFRYNTHR	TDNRKRSGNNFIRHTRHHIT	TARRHLIDGDGQRNIAFAQT			
m012-1	NIMFFQQA VD IRYFRHHTHR	TDNRKRSGSNFIRHTRHHIT	AARRHLIDGDGQRNIAFAQT			
	70	80	90	100	110	120
	130	140	150	160	170	180
a012-1.pep	PKLRSRQTVT VNHAARTFQSKQNL	IFRLGNQKRRRLMTQG	GFYGVCIQIAVKIQHKKAGF			
m012-1	XKLRSRQTVT VNHAARTFQSEQNL	IFRLGNQKRRRLMTQG	GFYGVCIQIAVKIQHKKAGF			
	130	140	150	160	170	180
	190	200	210	219		
a012-1.pep	LRFGRLPTLLQTLFLCFGRFLFLFLFLFLMFLCLFPAX					
m012-1	LRFGRLPTLLQTLFLCFGRFLFLFLFLFLMFLCLFPAX					
	190	200	210			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 87>:

g013.seq

```

1  aTgcctttga ccatgctgtg cagcaGGAcg tGCGGTTgt tcataatata
51  gtCcgaccGG AAAagcggAG GAAaCGCAGT GCCGCCTTCT TCCCCTTTCT
101 TGCCGTGGCA GGCGATGCag tTgGATTCTGT AACTTTTTTG CCCTTTtGtc

```

g013.pcp

```

1  MPLTLMCSRT  CGLFIIQSDR  KSGGNAVPRP  SPFLPWQAMQ  LDSYTFCPFV
51  MMLLSAAEAA  AQRQHKKMKAV  GSRVVFIGVS  PNVLKPCFLI  LPLRGEKFFW
101 PKSGI*

```

m013.seq

1	ATGCCTTTGA	CCATGCTGTG	CAGCAGCACC	TGCGGTTTTT	TCATGATGAA
51	GTCGGAGCGG	TAGAGCGGGC	GAAACATGGT	TCCGCGGCCCT	TCGCCCTTTT
101	TGCCCTGGCA	GCGCAGCGAG	TTGATTTCGT	ACACTTTTTG	CCCTTTTGTG
151	ATGATGCTGT	TGTCGGCGGC	AGAAGCGGCG	GCGCAGAAAGC	AGCCCAAGAC
201	GAGGGCGGTC	GGCAGTCGGG	TTGTGTTTCA	TGTTGTTTCC	TTCATGTTTG
251	AAACCTTGT	GTTGATTTTG	CGTAGCGGGT	GAAAGATTTT	TTTGCCGAAT
301	CAGTAG				

m013.pcp

1 MPLTLMCSST CGFFMMKSER XSGGNMVRPR SPFLPWQATQ LDSYTFCPFV  
51 MMLLSAAEAA AQKQPKTRAV GSRVVFIGVS FMEFETLLIL RSGXKIFLPN  
101 Q\*

a013.seq

1	ATGCCCTTTGA	CCATGCTGTG	CAGCAGCACC	TGCGGTTTTT	TCATGATGAA
51	GTCGAGCGCG	TAGAGCGGCG	GAAACATGTT	TCCGCGGCCCT	TCGCCCCCTTT
101	TGCCCTGGCA	GCGCAGCCAG	TTGGATTTCG	ACACTTTTTC	CCCTTTTGTG
151	ATGATGCTGT	TGTGCGCGCG	AGAAAGCGCG	GCGCAGAGGC	AGCCCCAAGAC
201	GAGGGCGGTC	GGCAGTCGGG	TTGTGTTTCAT	TGGTGTTCCT	TAAATGTTTG
251	AAACCTTGT	GTTGATTTTG	CGTAGCGGGT	GAAAGATTTT	CTTGCCGAAT
301	CGGTAG				

a013.pap

```

1  MPLTLMCSST CGFFMMKSER *SGGNMVP RP SPFLPWQATQ LDSYTFCPFV
51  MMLLSAAEAA AQRQPKTRAV GSRVVFIGVS LMFETLLLIL RSG*KIFLPN
101 R*

```

```

              10      20      30      40      50      60
m013.pep    MPLTMLCSSTCGFFMMKSERXSGGNMVRPRSPFLPWQATQLDSYTCFCFVMMLLSAAEAA
             |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a013         MPLTMLCSSTCGFFMMKSERXSGGNMVRPRSPFLPWQATQLDSYTCFCFVMMLLSAAEAA
              10      20      30      40      50      60

              70      80      90      100
m013.pep    AQKQPKTRAVGSRVVFIGVSLMFETLLLILRSGXKIFLPNQX
             ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a013         AQRQPKTRAVGSRVVFIGVSLMFETLLLILRSGXKIFLPNRX
              70      80      90      100

```

ORF 013 shows 73.3% identity over a 101 aa overlap with a predicted ORF (ORF 013.ng) from *N. gonorrhoeae*:

## m013/g013

	10	20	30	40	50	60
m013.pep	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTF	FCPFVMMLLSAAEAA				
g013	MPLTMLCSRTCGLFIIQSDRKSGGNAVPRPSPFLPWQAMQLDSYTF	FCPFVMMLLSAAEAA				
	10	20	30	40	50	60
	70	80	90	100		
m013.pep	AQKQPKTRAVGSRVVF	IGVSF-MFETLLLLLR-SG	KKIFLPNQX			
g013	AQRQHKKMAVGSRVVF	IGVSPNVLKP	CFLLPLRGEKFFWPKSGIX			
	70	80	90	100		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 93>:

```
g015.seq
1  ATGCAGTATC TGATTGTCAA ATACAGCCAT CAAATCTTCG TTACCATCAC
51  CATTTTGGTA TTCAACATCC GTTTTTTCCT ACTTTGGAAA AATCCAGAAA
101 AGCCCTTGGT CGGCTTTTGG AAAGCACTGC CCCACCTCAA CGACACGATG
151 CTGCTGTTTA CGGGATTGTG GCTGATGAAG ATTACCCATT TCTCCCGGTT
201 CAACGCGCCT TGGCTCGGCA CAAAAATCCT GCTCCTGTTT GCCTACATCG
251 CACTGGGCAT GGTAATGATG CGCGCCCGTC CGCGTTCGAC CAAGTTCTAC
301 ACCGTTTACC TGCTCGCTAT GTGTTGCATC GCCTGCATCG TTTACCTTGC
351 CAAAACCAAA GTCCTGCCAT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 94; ORF 015.ng>:

```
g015.pep
1  MQYLIVKYSH QIFVTITILV FNIRFFLLWK NPEKPLVGFW KALPHLNDTM
51  LLFTGLWLMK ITHFSPFNAP WLGTKILLLF AYIALGMVMM RARPRSTKFY
101 TVYLLAMCCI ACIVYLAKTK VLPF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 95>:

```
m015.seq (partial)
1  ..AAAATCAGAA AAGCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA
51  CGACACCATG CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT
101 TCTCCCGGTT CAACGCGCCT TGGCTCGGTA CAAAAATCCT GCTTCTGCTC
151 GCCTATATCG CATTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTCGAC
201 CAAGTTCTAC ACCGTTTACC TGCTCGCCAT GTGTTGCGTC GCCTGCATCG
251 TTTACCTTGC CAAAACCAAA GTCCTGCCTT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 96; ORF 015>:

```
m015.pep (partial)
1  ..KIRKALAGFW KALPHLNDTM LLFTGLWLMK ITHFSPFNAP WLGTKILLLL
51  AYIALGMMM RARPRSTKFY TVYLLAMCCV ACIVYLAKTK VLPF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 97>:

```
a015.seq
1  ATGCAGTATC TGATTGTCAA ATACAGCCAT CAAATCTTCG TTACCATCAC
51  CATTTTGGTA TTCAACATCC GTGTTTTCNT ACTTTGGAAA AATCCAGAAA
101 AGCCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA CGACACCATG
151 CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT TCTCCCGGTT
201 CAACGCGCCT TGGCTCGGTA CAAAAATCCT GCTTCTGCTC GCCTATATCG
251 CATTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTCGAC CAAGTTCTAC
301 ACCGTTTACC TGCTCGCCAT GTGTTGCGTC ACCTGCATCG TTTACCTTGC
351 CAAAACCAAA GTCCTGCCTT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 98; ORF 015.a>:

```
a015.pep
```

```

1  MQYLIVKYSH QIFVTITILV FNIRVFXLWK NPEKFLAGFW KALPHLNDTM
51  LLFTGLWLMK ITHFSPFNAP WLGTKILLLL AYIALGMMM RARPRSTKFY
101 TVYLLAMCCL TCIVYLAKTK VLPF*

```

m015/a015 96.7% identity over a 91 aa overlap

```

                                10      20      30
m015.pep                      KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                | | | | | | | | | | | | | | | | | |
a015      LIVKYSHQIFVTITILVFNIRVFXLWKNPEKFLAGFWKALPHLNDTMLLFTGLWLMKITH
              10      20      30      40      50      60

              40      50      60      70      80      90
m015.pep      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
              | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a015      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFYTVYLLAMCCLTCIVYLAKTKVLP
              70      80      90      100     110     120

m015.pep      FX
              ||
a015          FX

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 015 shows 94.5% identity over a 91 aa overlap with a predicted ORF (ORF 015.ng) from *N. gonorrhoeae*:

m015/g015

```

                                10      20      30
m015.pep                      KIRKALAGFWKALPHLNDTMLLFTGLWLMKITH
                                | | | | | | | | | | | | | | | | | |
g015      LIVKYSHQIFVTITILVFNIRFLLWKNPEKPLVGFWKALPHLNDTMLLFTGLWLMKITH
              10      20      30      40      50      60

              40      50      60      70      80      90
m015.pep      FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTKFYTVYLLAMCCVACIVYLAKTKVLP
              | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g015      FSPFNAPWLGTKILLLFAYIALGMVMRRARPRSTKFYTVYLLAMCCIIACIVYLAKTKVLP
              70      80      90      100     110     120

m015.pep      FX
              ||
g015          FX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 99>:

g018.seq

```

1  atGCAGCAGG GGCagttggt tggacgcgctc gcccgcaata AAGATATGCG
51  GAATgctggt CTGCATggtC AGCGGATCGG CAACGGGtac gccgcgcgcg
101 tctttgTCGA TATTGATGTT TTCCAAACCG ATATtgTCAA CGTTCGGACG
151 GCgACCTACG GCTGCCAACA TATATTGCGC AACAAATACG CCTTTTCGC
201 CATCCTGCTC CCAATGGACT tctACATTGC CGTCTSCGTC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT tctTCTCCGA ACACGGCTTT
301 CGCCTCGTCT GA

```

This corresponds to the amino acid sequence <SEQ ID 100; ORF 018.ng>:

g018.pep

```

1  MQQGQLVGRV ARNKDMRNAG LHGQRIGNGY AARVFVDIDV FQTDIVNVRT
51  ATYGCQHIFG NKYAFAFALL PMDFYIAVCV EFDLGFSIQM QFQFFESEHG

```

200

101 RLV\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 101>:

```
m018.seq
  1 ATGCAGCAGA GGCAGTTGGT TGGACGCATC GCCTGCGATG AAGATATGCG
 51 GAATACTGGT CTGCATGGTC AGCGGGTCGG CAACAGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATATT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC
251 TCGGTTTtag CATCCAGATG CAGTTTCAAT TCTTCGCCGA ACACGGCGTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 102; ORF 018>:

```
m018.pep
  1 MQQRQLVGRI ACDEDMRNTG LHGQRVGNRY AARIFFDIDI FQTDIVNVRT
 51 AAHGCGHIFG NKYAFAILL PMDFYIACV EFDLGFSIQM QFQFFAEHGV
101 RLV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 103>:

```
a018.seq
  1 ATGCAGCAGG GGCAGTTGGT TGGACGCGTC GCCCGCAATA AAGATATGCG
 51 GAATACTGGT CTGCATAGTC AGCGGATCGG CAACGSGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATGTT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCTACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCGTC GAGTTTGCC
251 TCGGTTTtag CATCCAAATG CAGTTTCAAT TCTTCACCGA ACACGGCTTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 104; ORF 018.a>:

```
a018.pep
  1 MQQGQLVGRV ARNKDMRNTG LHSQRIGNGY AARIFFDIDV FQTDIVNVRT
 51 AAYGCGHIFG NKYAFAILL PMDFYIACV EFDLGFSIQM QFQFFTEHGF
101 RLV*
```

m018/a018 86.4% identity over a 103 aa overlap

	10	20	30	40	50	60
m018.pep	MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCGHIFG					
	10	20	30	40	50	60
a018	MQQGQLVGRVARNKDMRNTGLHSQRIGNGYAARIFFDIDVFQTDIVNVRTAAYGCGHIFG					
	10	20	30	40	50	60
	70	80	90	100		
m018.pep	NKYAFAILLPMDFYIACVIEFDLGFSIQMQFQFFAEHGVRLVX					
	70	80	90	100		
a018	NKYAFAILLPMDFYIACVIEFDLGFSIQMQFQFFTEHGFRLVX					
	70	80	90	100		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 018 shows 84.5% identity over a 103 aa overlap with a predicted ORF (ORF 018.ng) from *N. gonorrhoeae*:

```
m018/g018
  1 MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCGHIFG
 51 AAHGCGHIFG NKYAFAILL PMDFYIACV EFDLGFSIQM QFQFFAEHGVRLVX
101 RLV*
  1 MQQGQLVGRVARNKDMRNTGLHSQRIGNGYAARVFDIDVFQTDIVNVRTATYGCQHIFG
 51 AAYGCGHIFG NKYAFAILL PMDFYIACV EFDLGFSIQM QFQFFTEHGFRLVX
101 RLV*
```



	70	80	90	100
m018.pep	NKYAFFAILLPMDFYIAVCIEFDLGFSIQMQFQFFAEHGVRLVX			
g018	NKYAFFAILLPMDFYIAVCIEFDLGFSIQMQFQFFSEHGFRVLX			
	70	80	90	100

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 105>:

```
g019.seq (partial)
1  ..ctgctggcgg ccctggtgct tgcgcgctgt tcttcgACAA ACaactTGCC
51  AGCCGGCAAG ACCCCGGCAG ACAATATAGA AactgcCgAC CTTTCGGCAA
101 GCGTTCCAC cgcCCTGCC GAACCGGAAG GAAAAACGCT GGCAGATTAC
151 GGCGGTACC CGTCCGCACT GGATGCAGTG AACAGAAC ACGATGCGGC
201 AGCCGCCGCC TATTTGGAAA Acgcaggaga cagCGcgatg gcGGAAAatg
251 tccgcaagga gtgGCTGa
```

This corresponds to the amino acid sequence <SEQ ID 106; ORF 019.ng>:

```
g019.pep (partial)
1  ..LLAALVLAAC SSTNTLPAGK TPAADNIETAD LSASVPTRPA EPEGKTLADY
51  GGYPSALDAV KQNNDAASAAA YLENAGDSAM AENVRKEWL*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 107>:

```
m019.seq (partial)
1  ATGTACCTAC CCTCTATGAA GCATTCCCTG CCGCTGCTGG CGGCCCTGGT
51  GCTTGCCGCG TGTTCTTCGA CAAACACACT GCCAGCCGCG AAGACCCCGG
101 CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
151 GCCGAACCCG AAAGAAAAAC GCTGGCAGAT TACGGCGGCT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTTGG GCGCACGCAG ACAGTGGACG CTGTTTGAC AGGAATACGC
351 CAAACTCGAA CCGGCAGGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCTG AACTGGTCAA AAATACGGGC
451 AAAGTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTGTTTGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGATT GGGCAGCCCG
601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCT TGGACTATTA
801 CGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCCTTGCGC GCCCGACGTT GGGACGAGCT GGCCTCCGTT
901 ATCTCGCATA TGCCCGAAAA ACTGCAAAAA AGCCCGACCT GGCTCTACTG
951 GCTGGCACGC AGCCGCGCCG CAACGGGCAA CACGCAAGAG GCGGAAAAAC
1001 TTTACAAACA GGCGGCAGCG ACGGGCAGGA ATTTTATGCG GGTGCTGGCA
1051 GGGGAAGAAT TGGGTGCGAA AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAAAC AGCGTCCGCC GCATGGCGGA AGACGGTGCA GTCAAACGCG
1151 CACTGGTACT GTTCCAAAAC AGCCAATCTG CCGGTGATGC AAAAATGCGC
1201 CGTCAGGCTC AGCGGGAATG GCGTTTGGCC ACACGCGGCT TTGACGAAGA
1251 CAAGCTGCTG ACCGCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACTCAA CTACACCTTG
1351 CGCTATATTT CGCCGTTTAA AGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCTTGGG TTTATGGGCT GATTGCTCAG GAAAGCCGCT
1451 TCGTTATAGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGAGGTT
1501 ATGCCTGCCA CCGCGCGCGA AATCGCGGCG AAAATCGGTA TGGATGCCCG
1551 ACAACTTTAC ACCGCCGACG GG...
```

This corresponds to the amino acid sequence <SEQ ID 108; ORF 019>:

```
m019.pep (partial)
1  MYLPSMKHSL PLLAALVLAA CSSTNTLPAG KTPADNIETA DLSASVPTRP
```

```

51  AEPERKTLAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARROWT LFAQEYAKLE PAGRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGV
251 GHYQSONLNV PAALDYYGKV ADRRQLTDDQ IEWYARAALR ARRWDELASV
301 ISHMPEKLQK SPTWLYWLAR SRAATGNTQE AEKLYKQAAA TGRNIFYAVLA
351 GEELGRKIDT RNNVPDAGKN SVRRMAEDGA VKRALVLFQN SQSAGDAKMR
401 RQAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNLYTL
451 RYISPFKDTV IRHAQNVNVD PAWVYGLIRQ ESRFVIGAQS RVGAQGLMQV
501 MPATAREIAG KIGMDAAQLY TADG...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 109>:

```

a019.seq
1  ATGTACCCAC CCTCTCTGAA GCATTCCCTG CCGTGCTGG TGGNCCTGGT
51  GCTTGCCGCG TGTTCCTNGA CAAACACACT GTCAGCCGAC AAGACCCCGG
101 CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCNGCCCT
151 GCCGAACCCG AANGAAAAAC GTNGGCAGA TACGGCGGCT ACCCGTCCCG
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCGTCGCC GCCTATTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTTGG GCGCGCGCAG ACAGTGGACG CTGTNTGCAC ANGAATATGC
351 NAAACTCGAA CCGGCANGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCGC AACTGGTCAA AAATACGGGC
451 AAAGTGCCTT CGGGCTGCAC CAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CTGCGCCGCT TGGACTATTA
801 NGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGC NNTNNGC NNNCGNNGT NGNANGANNT GGCNNCCGNN
901 ANCNCGNNNN TGCNNGANAA ACNNNNNNAN AGNCNNANNT NGNTNNANTG
951 NNTGGCACGC AGCCGCGCCG CNACGGGCAA CACGCAANAN GCGGANAAC
1001 TNTACAAACA GCGGCGAGCA NCGGGCANGA ATTTTATGCT NGTGCTGNCN
1051 GGGGAAGAGT TGGGCGGCAN AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAANC AGCGTCCTCC GTATGGCGGA AGACGGCGCG ATTAAGCGCG
1151 CGCTGGTGCT GTTCGGAAC AGCCGAACCG CCGCGGATGC GAAATGCGC
1201 CGTCNGGCTC AGCGCGAATG GCGTTTCGCC ACACGCGGCT TCGATGAAGA
1251 CAAGCTGAGT ACCGCGCGCG AAACCGCGTT CGACCACGGT TTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACCTCA CTACACCTTG
1351 CGCTACATTT CGNNNNNTNA NGACACGGTA ATCCGCCACG CGAAAATGT
1401 TAATGTCGAT CCGCGGTGGG TTTACGGGCT GATTCGTCAG GAAAGCCGCT
1451 TCGTTATGGG CGCGCAATCC CGCTAGGCG CGCAGGGGCT GATGAGGTT
1501 ATGCCTGCCA CCGCGCGCGA AATCGCCGCG AAAATCGGTA TGGATGCCCG
1551 ACAACTTTAC ACCGCGGACG GCAATATCCG TATGGGACG TGGTATATGG
1601 CGGACACCAA ACGCCGCTG CAAAACAACG AAGTCCTCGC CACCGCAGGC
1651 TATAACGCCG GTCCCGGACG GCGCGCCGGA TGGCAGGCGG ACACGCCCTT
1701 CGAAGGCGCG GTATATGCCG AAACCATCCC GTTTCCGAA ACGCGCGACT
1751 ATGTCAAAAA AGTGATGGCC AATGCCGCTC ACTACGCTC CCTCTTCGGC
1801 GCGCGCACCA TCCCGCTCAA ACAGCGTATG GGCATTGTCC CCGCCGCTG
1851 A

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 019.a>:

```

a019.pep
1  MYPPSLKHSLL PLLVXLVLAA CSXTNTLSAD KTPADNIETA DLSASVPTXP
51  AEFEXKTXAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARROWT LXAXEYAKLE PAXRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGV
251 GHYQSONLNV PAALDYYGKV ADRRQLTDDQ IEWYARAAXX XXXXXXXXXX
301 XXXXXXXXXX XXXXXXXXXAR SRAATGNTQX AXKLYKQAAA XGNIFYAVLX
351 GEELGRXIDT RNNVPDAGKX SVLRMAEDGA IKRALVLFN SETAGDAKMR
401 RXAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNLYTL

```

```

451 RYISXXXTDV IRHAQNVNVD PAWVYGLIRQ ESRFVMAQS RVGAQGLMQV
501 MPATAREIAG KIGMDAAQLY TADGNIRMGY WYMADTKRRL QNNEVLATAG
551 YNAGPGRARR WQADTPLEGA VYAETIPFSE TRDYVKKVMA NAAYVASLFG
601 APHIPLKQRM GIVPAR*

```

m019/a019 88.9% identity over a 524 aa overlap

	10	20	30	40	50	60
m019.pep	MYLPSMKHSLPLLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAPERKTLAD					
a019	MYPPSLKHSLPLLVXLVLAACSXTNTLSADKTPADNIETADLSASVPTXPAPPEKTXAD					
	10	20	30	40	50	60
m019.pep	70	80	90	100	110	120
	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKS LGARRQWTLFAQEYAKLE					
a019	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKS LGARRQWTLXAXEYAKLE					
	70	80	90	100	110	120
m019.pep	130	140	150	160	170	180
	PAGRAQEVECYADSSRNDYTRAAELVKNTGKLP SGCTK LLEQAAASGLLDGNDAWRRVRG					
a019	PAXRAQEVECYADSSRNDYTRAAELVKNTGKLP SGCTK LLEQAAASGLLDGNDAWRRVRG					
	130	140	150	160	170	180
m019.pep	190	200	210	220	230	240
	LLAGRQTDDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEME SGLSL					
a019	LLAGRQTDDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEME SGLSL					
	190	200	210	220	230	240
m019.pep	250	260	270	280	290	300
	EQRSFAWGVLGHYQS QNLNVPAALDYXGKVADRRQLTDDQIEWYARAALRARRWDELASV					
a019	EQRSFAWGVLGHYQS QNLNVPAALDYXGKVADRRQLTDDQIEWYARAAXXRXRXAXX					
	250	260	270	280	290	300
m019.pep	310	320	330	340	350	360
	ISHMPEKLQKSPTWLYWLARSRAATGNTQEAELKYQAAATGRNFYAVLAGEELGRKIDT					
a019	XXXXXXXXXXXXXXXXXARSRAATGNTQXAKLYQAAAGXNFXAVLXGEELGRXIDT					
	310	320	330	340	350	360
m019.pep	370	380	390	400	410	420
	RNNVPDAGKNSVRRMAEDGAVKRALVLFQNSQSAGDAKMRRQAQAEWR FATRGFDEDKLL					
a019	RNNVPDAGKXSVLRMAEDGAIKRALVLFNRNSRTAGDAKMRRXAQAEWR FATRGFDEDKLL					
	370	380	390	400	410	420
m019.pep	430	440	450	460	470	480
	TAAQTAFDHGFYDMVNSAERTDRKLNITLRYISPFKDTVIRHAQNVNVDPAWVYGLIRQ					
a019	TAAQTAFDHGFYDMVNSAERTDRKLNITLRYISXXXTVIRHAQNVNVDPAWVYGLIRQ					
	430	440	450	460	470	480
m019.pep	490	500	510	520		
	ESRFVIGAQS RVGAQGLMQVMPATAREIAGKIGMDAAQLYTADG					
a019	ESRFVMAQS RVGAQGLMQVMPATAREIAGKIGMDAAQLYTADGNIRMGY WYMADTKRRL					
	490	500	510	520	530	540
a019	QNNEVLATAGYNAGPGRARRWQADTPLEGAVYAETIPFSETRDYVKKVMANAAYVASLFG					
	550	560	570	580	590	600

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 019 shows 95.5% identity over a 89 aa overlap with a predicted ORF (ORF 019.ng) from *N. gonorrhoeae*:

```

g019/m019

          10      20      30      40      49
g019.pep      LLAALVLAACSSNTNLPAGKTPADNIETADLSASVPTRPAEPEGKTLAD
              |||||
m019          MYLPSMKHSLPLLAALVLAACSSNTNLPAGKTPADNIETADLSASVPTRPAEPEKTLAD
              10      20      30      40      50      60

          50      60      70      80      89
g019.pep      YGGYPSALDAVKQNNDAAAAYLENAGDSAMAENVRKEWL
              |||||
m019          YGGYPSALDAVKQNDAAVAAYIENAGDSAMAENVRNEWLKSLGARRQWTLFAQEYAKLE
              70      80      90      100     110     120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 111>:

```

g023.seq
1   ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GCGATTGGGT
51  AATGCAGCGT GCGACTGCGG TTATTATGTT GATTTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTGCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
151 TTTAGTCAAG CTTGGGTAAA AGTATTTACC CAAGTGAGCT TTATCGCCGT
201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCCTCGG CGTGCGTTTG TTTTGCAGG TTGCCACCAT TGcttGGCTG
301 GTCGGCTGCC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 023.ng>:

```

g023.pep
1   MVERKLGAH YGLRDWVMQR ATAVIMLIYT VALLVLFAL PKEYPAWQAF
51  FSQAWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

```

m023.seq
1   ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GCGATTGGGT
51  GATGCAACGT GCGACTGCGG TTATTATGTT GATTTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTCCCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT
201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCCTCGG CGTGCGTTTG TTTTGCAGG TTGCCACCAT CGTTTGGCTG
301 GTCGGCTGTC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 023>:

```

m023.pep
1   MVERKLGAH YGLRDWVMQR ATAVIMLIYT VALLVLFSL PKEYSAWQAF
51  FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

```

a023.seq
1   ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GGGATTGGGC
51  GATGCAACGT GCGACCGCGG TTATTATGTT GATTTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTGCTCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT

```

205

```

201 ATTCTTGAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATNA
251 AACCTTCGG CGTGCGTTTG TTTTTCAGG TTGCCACCAT CGTCTGGCTG
301 GTCGGCTGCT TGGTGTATTC AATTAAAGTA ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 023.a>:

```

a023.pep
1 MVERKLTGAH YGLRDWAMQR ATAVIMLIYT VALLVVLFA L_PKEYSAWQAF
51 FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYXKPFGRV LFLQVATIVWL
101 VGCLVYSIKV IWG*

```

m023/a023 96.5% identity over a 113 aa overlap

```

          10      20      30      40      50      60
m023.pep MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFSLPKEYSAWQAFFSQTWVKVFT
          |||||
a023      MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFALPKEYSAWQAFFSQTWVKVFT
          10      20      30      40      50      60

          70      80      90     100     110
m023.pep QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          |||||
a023      QVSFIAVFLHAWVGIRDLWMDYXKPFGRVRLFLQVATIVWLVGCLVYSIKVIWGX
          70      80      90     100     110

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 023 shows 97.3% identity over a 113 aa overlap with a predicted ORF (ORF 023.ng) from *N. gonorrhoeae*:

```

g023/m023
          10      20      30      40      50      60
g023.pep MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFALPKEYPAWQAFFSQAWVKVFT
          |||||
m023      MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFSLPKEYSAWQAFFSQTWVKVFT
          10      20      30      40      50      60

          70      80      90     100     110
g023.pep QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          |||||
m023      QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          70      80      90     100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 117>:

```

g025.seq
1 ATGTTGAAAC AAacgACACT TTTGGCAGCT TGTACCGCCG TTGCCGCTCT
51 GTTGGGCGGT TGcgCCACCC AACAGCCTGC TccTGTcATT GCAGGCAATT
101 CAGGTATGCA GACCGTATCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
151 ACGCCGTACA ATGCCGCTCC TGCCGCCAac gatgcGCCgT ATGTGCCGCC
201 CGTGCAAact gcgcgggttT ATTGCCTCC TGCTTATGTT CCGCcgctCTG
251 CACCTGCCGT TTCGGgtaca tatgtTCCTT CTTACGCACC CgtcgACATC
301 aacgCGGCGa cgCataCTAT TGTGCGTGGC GACACgGtgt acaACATTc
351 caaAcgCtac CATATCTCTC AAGACGATTT CCGTGCCTGG AACGGCATGA
401 CCGACAATAC GTTGAGCATC CgTCAGATTG TTAAAGTCAA ACCGGCaggA
451 TATGCCGCAC CGAAAACCGC AGCCGTAGAA AGCAGGCCCG CCGTACCGGC
501 TGCCGCGCAA ACCCTGTGA AACCGCCGC gcaACCGCCC GTTCAGTCCG
551 CGCCCAACC TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCCCCC
601 GCGCCGCCC CGCAATCTCC TGCCGCTTCG CCTTCGGCA CGCGTTCGGT
651 CGCGGCATT GTTTGCAGC GTCCACCCA AGGTAAAGTG GTTGCCGATT

```

```

701 TCGGCGGCGG CAACAAGGGT GTCGATATTG CCGGCAATGC CGGACAACCC
751 GTTTTGGCGG CGGCTGACGG CAAAGTGGTT TATGCCGTT CAGGTTTGAG
801 GGGATACGGA AACTTGGTCA TCATCCAGCA CAATTCCTCT TTCCTGACCG
851 CGTACGGGCA CAACCAAAA TTGCTGGTCC GCGAAGGTCA GCAGGTCAAA
901 CGCJGTACGC AGGTTGCTTT GATGGGTAAT ACCGATGCTT CCAGAACGCA
951 GCTTCATTTC GAGGTGCGTC AAAACGGCAA ACCGGTTAAC CCGAACAGCT
1001 ATATCGCGTT CTGA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 025.ng>:

```

g025.pep
1 MLKQTTLLAA CTAVAALLGG CATQQPAPVI AGNSGMQTVS SAPVYNPYGA
51 TPYNAPAAAN DAPVVPVQT APVYSPPAYV PPSAPAVSGT YVPSYAPVDI
101 NAATHITVRG DTVYNISKRY HISQDDFRW NGMTDNTLSI GQIVKVKPAG
151 YAAPKTAAVE SRPAVPAAAO TPVKPAAQPP VQSAPQPAAP AAENKAVPAP
201 APAPQSPAAS PSGTRSVGGI VWQRPTQKQV VADFGGNGK VDIAGNAGQP
251 VLAAADGKVV YAGSGLRGY NLVIIQHNSS FLTAYGHNQK LLVGEQQQVK
301 RGQQVALMGN TDASRTQLHF EVRQNGKPVN PNSYIAF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```

m025.seq (partial)
1 ..GTGCCGCGG TGCAAAGCGC GCCGGTTTAT ACGCCTCCTG CTTATGTTCC
51 GCCGTCTGCA CCTGCCGTTT CGGGTACATA CGTTCCTTCT TACGCACCCG
101 TCGACATCAA CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC
151 AACATTTCOA AACGCTACCA TATCTCTCAA GACGATTTC GTGCGTGGAA
201 CGGCATGACC GACAATACGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC
251 CGGCAGGATA TGCCGCACCG AAAGCCGCAG CCGTAAAAAG CAGGCCCGCC
301 GTACCGGCTG CCGCGCAACC GCCCGTACAG TCCGCACCCG TCGACATTAA
351 CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC AACATTTCCA
401 AACGCTACCA TATCTCTCAA GACGATTTC GTGCGTGGAA CGGCATGACC
451 GACAATATGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC CGGCAGGATA
501 TGCCGCACCG AAAACCGCAG CCGTAGAAAG CAGGCCCGCC GTACCGGCTG
551 CCGTGCAAAC CCCTGTGAAA CCGCGCGCGC AACCGCCTGT GCAGTCCGCG
601 CCGCAACCTG CCGCGCCCGC TGCGGAAAAT AAAGCGGTTT CCGCGCCCGC
651 CCGCAATCT CCTGCCGCTT CGCCTTCGCG CACGCGTTCC GTGCGGCGCA
701 TTGTTTGGCA GCGTCCGACG CAAGGTAAAG TGTTTGCCGA TTTCGCGCGC
751 AACAACAAGG GTGTCGATAT TGCCGGTAAT GCGGGACAGC CCGTTTGGC
801 GGCGGCTGAC GGCAAAAGTGG TTTATGCCGG TTCAGGTTT AGGGGATACG
851 GAAACTTGGT CATCATCCAG CATAATTCTT CTTTCCTGAC CGCATACGGG
901 CACAACCAAA AATTGCTGGT CGGCGAGGGG CAGCAGGTCA AACGCGGTCA
951 GCAGGTTGCT TTGATGGGCA ATACCGATGC TTCCAGAACG CAGCTTCATT
1001 TCGAGGTGCG TCAAAACGGC AAACCGGTTA ACCCGAACAG CTATATCGCG
1051 TTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 025>:

```

m025.pep (partial)
1 ..VPPVQSAPVY TPPAYVPPSA PAVSGTYVPS YAPVDINAAT HTIVRGDTVY
51 NISKRYHISQ DDFRAWNGMT DNTLSIGQIV KVKPAGYAAP KAAAVKSRPA
101 VPAAAQPPVQ SAPVDINAAT HTIVRGDTVY NISKRYHISQ DDFRAWNGMT
151 DNMLSIGQIV KVKPAGYAAP KTAAVESRPA VPAAVQTPVK PAAQPPVQSA
201 PQPAAPAAEN KAVPAPAPQS PAASPSGTRS VGGIVWQRPT QGKVVADFGG
251 NNKGVDIAGN AGQPVLAAD GKVVYAGSGL RGYGNLVIIQ HNSSLTAYG
301 HNQKLLVGEQ QQVKRGQQA LMGNTDASRT QLHFEVRQNG KPVNPNNSYIA
351 F*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

a025.seq
1 ATGTTGACAC CAACAACACT TTAGGTAGCT TGTACCGCCC TTGCCGCTCA
51 GTTGGGCGGA TGCCCCACCC AACACCTTC TCCTGTCAAT GCAGGCAATT
101 CAGGTATGCA GACCGTACCG TCTGCGCCGG TTTACAATCC TTATGGCGCA

```

207

```

151  ACGCCGTACA ATGCCGCTCC TGCCGCCAAC GATGCGCCGT ATGTGCGCGC
201  GGTGCAAAGC GCGCCGGTTT ATANGCCTCC TGCTTATGTT CCGCCGTCTG
251  CACCTGCCGT TTCGGGTACA TACGTTCTCT CTTACGCANC CGTCGACATC
301  AACCGCGCGA CCCATACTAT TGTGCGCGGC GACACCGTGT ACAAGATTTT
351  CAAATGCTAC CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
401  CCGACAATAC GTTGAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCAGGA
451  TATGCCGCAC CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC
501  TGCCGCGCAA CCGCTCGTAC AGTCCGCACC CGTCGACATC AACGCGCGCA
551  CGCATACTAT TGTGCGCGGC GACACGGTGT ACAACATTTT CAAACGCTAC
601  CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA CCGACAATAC
651  GTTGAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCAGGA TATGCCGCAC
701  CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC TGCCGTGCAA
751  ACCCCTGTGA AACCCGCCGC GCAACCGCCT GTGCAGTCCG CGCCGCAACC
801  TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCGCCC GCCCCGCAAT
851  CTCCTGCCGC TTCGCCTTCC GGCACGCGTT CCGTCGCGCG CATTGTTTGG
901  CAGCGTCCGA CGCAAGGTAA AGTGGTTGCC GATTTCGCGC GCAACAACAA
951  GGGTGTTCGAT ATTGCAGGAA ATGCGGGACA GCCCGTTTGG GCGGCGGCTG
1001 ACGGCAAAGT GGTTTATGCA GGTTCGCGTT TGAGGGGATA CGGCAATTG
1051 GTCATCATCC AGCATAATTC TTCCTTCTTG ACCGCATACG GGCACAACCA
1101 AAAATTGCTG GTCGGCGAAG GCCAGCAGGT CAAACGCGGG CAGCAGGTCG
1151 CTTTGATGGG CAATACCGAG GCTTCTAGAA CGCAGCTTCA TTTCGAGGTG
1201 CGGCAAAACG GCAAACCGGT TAATCCGAAC AGCTATATCG CGTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 025.a>:

```

a025.pep
1  MLTPTTL*VA CTALAAQLGG CPTQHPSPVI AGNSGMQTVP SAPVYNPYGA
51  TPYNAAPAA DAPYVPPVQS APVYXPPAYV PPSAPAVSGT YVPSYAXVDI
101 NAATHTIVRG DTVYKISKCY HISQDDFRW NGMTDNTLSI GQIVKVKPAG
151 YAAPKAAAVK SRPAVPAAAQ PLVQSAPVDI NAATHTIVRG DTVYNISKRY
201 HISQDDFRW NGMTDNTLSI GQIVKVKPAG YAAPKAAAVK SRPAVPAAVQ
251 TPVYKAAQPP VQSAPQPAAP AAENKAVPAP APQSPAASPS GTRSVGGIVW
301 QRPTQGVVVA DFGGNNKGVD IAGNAGQPVL AAADGKVVA GSGLRGYGNL
351 VIIQHNSFL TAYGHNQKLL VEGEQQVVRG QQVALMGNT EASRTQLHFEV
401 RQNGKPVNPN SYIAF*

m025/a025 97.4% identity over a 351 aa overlap

                                10      20      30
m025.pep                                VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
                                |||:|||||:|||||:|||||:|||||:|||||:
a025      GMQTVPSAPVYNPYGATPYNAAPAANDAPYVPPVQSAPVYXPPAYVPPSAPAVSGTYVPS
                                40      50      60      70      80      90

                                40      50      60      70      80      90
m025.pep      YAPVDINAATHTIVRGDTVYNISKRYHISQDDFRWNGMTDNTLSIGQIVKVKPAGYAAP
                || |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a025      YAXVDINAATHTIVRGDTVYKISKCYHISQDDFRWNGMTDNTLSIGQIVKVKPAGYAAP
                100      110      120      130      140      150

                                100      110      120      130      140      150
m025.pep      KAAAVKSRPAVPAAPVQVQSAPVDINAATHTIVRGDTVYNISKRYHISQDDFRWNGMT
                |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a025      KAAAVKSRPAVPAAPVQVQSAPVDINAATHTIVRGDTVYNISKRYHISQDDFRWNGMT
                160      170      180      190      200      210

                                160      170      180      190      200      210
m025.pep      DNMLSIGQIVKVKPAGYAAPKTAAVESRPVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
                || |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a025      DNTLSIGQIVKVKPAGYAAPKAAAVKSRPAVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
                220      230      240      250      260      270

                                220      230      240      250      260      270
m025.pep      KAVPAPAPQSPAAASPSGTRSVGGIVWQRPTQGVVADFGGNNKGVDIAGNAGQPVLAAAD
                |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

208

```

a025      KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAD
           280      290      300      310      320      330

m025.pep      280      290      300      310      320      330
GKVVYAGSGLRGYGNLVIIQHNSSELTAYGHNQKLLVGEGQQVKRGQQVALMGNTDASRT
|||||
a025      GKVVYAGSGLRGYGNLVIIQHNSSELTAYGHNQKLLVGEGQQVKRGQQVALMGNTEASRT
           340      350      360      370      380      390

m025.pep      340      350
QLHFEVRQNGKPVNPNSYIAFX
|||||
a025      QLHFEVRQNGKPVNPNSYIAFX
           400      410

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 025 shows 75.6% identity over a 353 aa overlap with a predicted ORF (ORF 025.ng) from *N. gonorrhoeae*:

m025/g025

```

m025.pep      10      20      30
VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
|||||
g025      GMQTVSSAPVYNPYGATPYNAAPAANDAPYVPPVQTAPVYSPPAYVPPSAPAVSGTYVPS
           40      50      60      70      80      90

m025.pep      40      50      60      70      80      90
YAPVDINAATHITVRGDTVYNISKRYHISQDDFAWNGMTDNTLSIGQIVKVKPAGYAAP
|||||
g025      YAPVDINAATHITVRGDTVYNISKRYHISQDDFAWNGMTDNTLSIGQIVKVKPAGYAAP
           100     110     120     130     140     150

m025.pep      100     110     120     130     140     150
KAAAVKSRPAVPAAPAAQPPVQSAPVDINAATHITVRGDTVYNISKRYHISQDDFAWNGMT
|
g025      K-----

m025.pep      160     170     180     190     200     210
DNMLSIGQIVKVKPAGYAAPKTAAVESRPAPVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
|||||
g025      -----TAAVESRPAPVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
           160     170     180     190

m025.pep      220     230     240     250     260
KAVPAPAP--QSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAD
|||||
g025      KAVPAPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAD
           200     210     220     230     240     250

m025.pep      270     280     290     300     310     320
ADGKVVYAGSGLRGYGNLVIIQHNSSELTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
|||||
g025      ADGKVVYAGSGLRGYGNLVIIQHNSSELTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS
           260     270     280     290     300     310

m025.pep      330     340     350
RTQLHFEVRQNGKPVNPNSYIAFX

```



209

g025                    |||||  
                       RTQLHFEVRQNGKPVNPNSYIAFX  
                       320                    330

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 113>:

g031.seq  
 1 ATGGTGTCCC TCCGCTTCAG ATTCCGCAAC CACTTTAAAC GCCGACATTC  
 51 TGACAATTTC CTTTCCGCC AGCCAAATAT CATGCGTATC TTTCGGTTCCG  
 101 GGCTTGTGG GCATGGCAAC CTTCAACAGC CGCGCCATCA CAGGAATCGT  
 151 CGTTCCCTGA ATCAGCAGCG ACAGCACCAC CACGGCAAAC GCCACATCAA  
 201 ACAGCAGGTG CGATTGGGA ACGCCCATCA CCAGCGGCAT CATCGCCAGC  
 251 GAAATCGGTA CGGCTCCTCG CAAGCCCAAC CAACTGATAT ACGCCTTTTC  
 301 ACGCAGGCTG TAATTGAATT TCCACAAACC GCCGAACACT GCCAGCGGAC  
 351 GCGCGACCAG CATCAGGAAC GCGCAATCG CCAAGGCTTC CGCCGCCCTG  
 401 TCCAACACGC CGCGGGGAGA AACCAGCAGA CCGAGCATGA CGAACAAAGT  
 451 TGCCTGCGCC AGCCAAGCCA AACCGTCCAT CACACGCAA ACGTGTTCGG  
 501 TcgcACGGTT GCGCTGGTTA CCGACAATGA TGCCGGCAAG GTAAACCGCC  
 551 AAAAAAGCCG TGCCGCCTAT GGTATTGGTA AACGCAAACA CAAGCAGCCC  
 601 GCCCGACACA ATCATCAGCG CGTACAGACC TTCGGTACAC acctccaatt  
 651 cccaatcaac gtcataagctg tctcccggtg taaaatgttc ttcaactcag  
 701 aatccccccc ttcttccag cccgaaacct tcatgtgtta naccctgggg  
 751 tgccccaacg gatttagtaa cctcccaatg actctgcttg tcgccccctt  
 801 cgcccgcttt ctccctccg gaaaacttgt tgtccccgtc ttacattaa

This corresponds to the amino acid sequence <SEQ ID 114; ORF 031.ng>:

g031.pep  
 1 MVSLRFREFG HFKRRHSDNF LFRQPNIMRI FRFGLVGHGN LQQPRHHRNR  
 51 RSLNQQRQHH HGKRRHIKQOV RIGNAHHQRH HRQRNRYGSS QAQPTDIRLF  
 101 TQAVIEFPQT AEHCQRTDQ HQERRNRQGF RRPVQHAGGR NQOTEHDEQS  
 151 CLRQPSQTVH HTQNVFRRTV ALVTDNDAGK VNRQKAAAY GIGKRKHQKQ  
 201 ARHNNHQRVQT FRTHLQFPIN VIAVSRVKMF FTSESPSSQ PETFMCXTLG  
 251 CPNGFSNLPM TLLVAPFARF LLPGKLVPV LH\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

m031.seq (partial)  
 1 ...CGCCTGAAGC ACGGTGTCGG ACTGCATTTC TATTCGGCTA TACGCCTTTT  
 51 CACGCAGGCT GTAATTGAAT TTCCACAAAC CGCCGAACAC TGCCGACGGA  
 101 CGCGCGACCA GCATCAGGAA CGCCGCAATC GCCAAGGCTT CGCCGCCCTT  
 151 GTCCAACACG TTGGCAGGAG AAACCAGCAG CAAAGGCATT CCCAAACGTG  
 201 CGGACAAAGT GGTGAAACC ACGCTCAGAA ACAACAGTGC GCCACCCGGC  
 251 AG....

This corresponds to the amino acid sequence <SEQ ID 116; ORF 031>:

m031.pep (partial)  
 1 ...RLKHGVLHF YSAIRLFTQA VIEFPQTAEH CRRTDQHQE RNRQGFRRP  
 51 VQHVGRNRNQ QRHSQTCGQS GRNHAQKQQC ATRQ....

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

a031.seq  
 1 ATACGCCTTT TCACGCAGGC TGTAATTGAA TTTCCACAAA CCGCCGAACA  
 51 CTGCCGCGCG ACGCGCGACC AGCATCAGGA ACGCCGCAAT CGCCAAGGCT  
 101 TCCGCCGCCC CGTCCAACAC GTTGGCAGGA GAAACCAGCA GCAAAGGCAT  
 151 TCCCAAACGT GCGGACAAAG TGGTCGAAAC CACGCTCAGA AACAAACAGT  
 201 CGCCACCCGG CAG

This corresponds to the amino acid sequence <SEQ ID 118; ORF 031.a>:

a031.pep (partial)  
 1 IRLFTQAVIE FPQTAEHCRR TRDQHQERRN RQGFRRPVQH VGRNRQQQRH  
 51 SQTGQSGRN HAQKQCCATR Q

```

      10      20      30      40      50      60
m031.pep  RLKHGVGLHFYSAIRLFTQAVIEFPQTAEHCRRTDQHQERRNRQGFRRPVQHVGRNRQQ
          |||
a031      IRLFTQAVIEFPQTAEHCRRTDQHQERRNRQGFRRPVQHVGRNRQQ
          10      20      30      40

      70      80
m031.pep  QRHSQTCGQSGRNHAQKQQCATRQ
          |||
a031      QRHSQTCGQSGRNHAQKQQCATRQ
          50      60      70

```

ORF 031 shows 60.0% identity over a 85 aa overlap with a predicted ORF (ORF\_031.ng) from *N. gonorrhoeae*:

m031/g031

[illegible]

g032.seq

1	ATGCGGCGAA	ACGTGCCTGC	CGTCGCCGTA	TTGCGCCGCC	CACGATTCGA
51	GGCGTTTTTG	GATTTGGCGT	TGGCTCAGGC	GCGTGCCGTT	CCTGCCGGTA
101	AACAGGGCTT	TGCCGTCGCA	TGCCGCTGTA	CGCAGCGGCA	GATAGTTTT
151	CAAGGCTTCC	ACGCTTTTGC	CGGTGACCGG	AACCTGACGC	TGCTTGCGCC
201	CTTTGCCCGT	AACGTGTACC	CACGCTTCGT	CCAAATATAC	ATCATCTGCA
251	TTCAGACCGT	GTATCTCGCT	CACGCGCAAA	CCGCTGCCGT	ACATCAGCTC
301	GAACAGCGCG	TGGTCGCGCA	CGCCAGCGCG	GTCCGCCCGC	TCCACGGGCA
351	AATCCAACAT	CCGGTTCAGC	CATTCTCGCG	GCAGGGCTTT	TGGTACGCGC
401	TCGGGCTGCT	TCGGCGGTTT	GATGTCGGCG	GTCCGGTCGG	CGCGCATCAG
451	CCGCGCTTTG	ACCAGCCAGG	CGCAATACTG	CCGCCACGCC	GACGACTTGC
501	GCGCCAGCTT	CCGACCGCTC	AAACCGCGCT	GCACAGACCG	CTCGAACCGC
551	GcgTAAAT	CGCGCCGCGA	CAAGTCTGCG	GGCAGCGcg	ctgcaTCTTC
601	AGACGGCATT	TGTGCCAACA	GTGCAAAACG	TTCTTCCAAA	TGCGCCCGGT
651	ATGCCGCAAC	CGTGTGCTCT	GACTTGCCCT	CGCGCACGAT	GTTTTCACAA
701	TAAAGCTCAA	AATacgcgc	AAACcgcTCC	AAAAcCATAA	CCGTCCCCACA
751	CAAAATATCAA	AAAACCATGT	A		

g032.pcp

1 MRRNVPAAVAV LRRPRFEAFI DLALAQARAV PAGKQGFVAV CRLTQRQIVF  
51 QGFHAFAGQR NLTLAPFAG NVYPREVQIY IICIQAVYLA HAQTAHVHQL

```

101 EQRVVAHRQR VAAVHGQIQH PVQPFRLRQGF GYALGLLRRF DVGGRVGAHQ
151 PAFDQPGAIL PPRQLARQR PTQTALRQP PQRRRKIAPR QVLRHAACIF
201 RRHLCQQCKQ FFQIAPVCRN RVLRLALAH DVFQISVKIRR KPVQNHNRPT
251 QISKNQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

m032.seq (partial)
1 ATGCGGCGAA ACGTGCmTGC mGTGCGCGTT kTGCGCCGCC CATTGCGCCA
51 AACGTTTTTG GATTGGCGT TGGCTCAGGC GCGTGCCGT CCTGCCGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA TGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGACCAGCGG CACCTGCCGC TgTT.GCGCC
201 CTTTGCCGAT AACGTGTACC CACGCTCGT CCAAATAGAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAA CCGCTGCCGT ACATCAGTTC
301 GAACAGGGCG TGGTCGCGCA CGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTTACC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG CGTGCACTCAG
451 GCCGCGCTTT ACCAGCCAAA CGCAATACTG CCGCCAAGAC GAAAGCTTGC
501 GAGCCAGCGT CCGTTCCCCC AAACCGCG...

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 032>:

```

m032.pep (partial)
1 MRRNVXAVAV XRRPLRQTFI DLALAQARAV PAGKQGFVAV CRLTQRQIVF
51 QGFHAFADQR HLPLXAPFAD NVYPRXVQID IICIQAVYLA HAQTAAVHQF
101 EQGVVAHRQR VAAVHGQIQH PVQPFRLRQGF GYALGLLRRF DVGGRVGVHQ
151 AALYQPNAIL PPRKRLASQR PFPQTA...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

```

a032.seq
1 ATGCGGCGAA ACGTGCCTGC CGTCGCCGT TTGCGCCGCC CATTGCGCCA
51 AACGTTTTTG GATTGGCGT TGGCTCAGGC GCGTGCCGT CCTGCCGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGCCGC TGCTTGCGTC
201 CTTTGCCGGT AACGTGTACC CACGCCTCGT CCAAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAA CCGCTGCCGT ACATCAGTTC
301 GAACAGCGCG TGATCGCGCA CGCCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTTACG CATTCCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG TATGCAGCAG
451 ACCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCAAGAC GACAGCTTGC
501 GCGCCAGCGT CCGCGCATTC AAACCGCGT GCGACAGCCG CCGCAACGCC
551 CCGGTAAAAT CGCGCTGCGA CAAGCCCTGC GGCACGCCGC CTGCATCTTC
601 AGACGGCATT TGTGCCAACA GCGCAAACAG TTCTTCCAAA TCGCGCGGCT
651 ATGCCGCCAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701 TAAGCGTCAA AATGCGCCGC AAACCGTCC AAAACCATAA CCGCCCCACA
751 CAAATATCAA AAAACAGTG A

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 032.a>:

```

a032.pep
1 MRRNVPAAVAV LRRPLRQTFI DLALAQARAV PAGKQGFVAV CRLTQRQIVF
51 QGFHAFAGQR NLPLASFAG NVYPRLVQIY IICIQAVYLA HAQTAAVHQF
101 EQRVIAHRQR VAAVHGQIQH PVQPFRLRQGF GYALGLLRRF DVGGRVGMQQ
151 TAFDQPGAIL PPRQLARQR PRIQTALRQP PQRRRKIALR QALRHAACIF
201 RRHLCQQRKQ FFQIAPVCRH RVLRLALAH DVFQISVKMRR KPVQNHNRPT
251 QISKKQ*

```

m032/a032 88.1% identity over a 176 aa overlap

```

          10      20      30      40      50      60
m032.pep  MRRNVXAVAVXRRPLRQTFIDLALAQARAVPAGKQGFVAVRCRLTQRQIVFQGFHAFADQR
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a032      MRRNVPAAVAVLRRPLRQTFIDLALAQARAVPAGKQGFVAVRCRLTQRQIVFQGFHAFAGQR
          10      20      30      40      50      60

```

212

	70	80	90	100	110	120
m032.pep	HLPLXAPFADNVYPRXVQIDIICIQAVYLAHAQTAAVHQFEQGVVAHRQRVAAVHGGQIQH					
a032	NLPLLASFAGNVYPRLVQIYIICIQAVYLAHAQTAAVHQFEQRVIAHRQRVAAVHGGQIQH					
	70	80	90	100	110	120
	130	140	150	160	170	
m032.pep	PVQPFLLRQGFYALGLLRRFDVGGRVGVHQAALYQPNAILPPRRKLASQRPFPQTA					
a032	PVQPFLLRQGFYALGLLRRFDVGGRVGMQQTAFDQPGAILPPRRQLARQRPRIQTALRQP					
	130	140	150	160	170	180
a032	PQRRRKIALRQALRHAACIFRRHLCQQRKQFFQIAPVCRHRVLRLLALAHDFQISVKMRR					
	190	200	210	220	230	240

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 032 shows 86.4% identity over a 176 aa overlap with a predicted ORF (ORF 032.ng) from *N. gonorrhoeae*:

m032/g032

	10	20	30	40	50	60
m032.pep	MRRNVXAVAVXRRPLRQTFDLALAQARAVPAGKQGFVRCRLTQRQIVFQGFHAFADQR					
g032	MRRNVPAVAVLRRPRFEAFDLALAQARAVPAGKQGFVRCRLTQRQIVFQGFHAFAGQR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m032.pep	HLPLXAPFADNVYPRXVQIDIICIQAVYLAHAQTAAVHQFEQGVVAHRQRVAAVHGGQIQH					
g032	NLTLLAPFAGNVYPRFVQIYIICIQAVYLAHAQTAAVHQFEQRVVAHRQRVAAVHGGQIQH					
	70	80	90	100	110	120
	130	140	150	160	170	
m032.pep	PVQPFLLRQGFYALGLLRRFDVGGRVGVHQAALYQPNAILPPRRKLASQRPFPQTA					
g032	PVQPFLLRQGFYALGLLRRFDVGGRVGAHQPAFDQPGAILPPRRQLARQRPVTQTALRQP					
	130	140	150	160	170	180
g032	PQRRRKIAPRQVLRHAACIFRRHLCQQCKQFFQIAPVCRNRVLRLLALAHDFQISVKIRR					
	190	200	210	220	230	240

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 115>:

g033.seq

```

1  ATGGCGGCGG CGGACAAACT CTTGGGCGGC GACCGCCGCA GCGTCGCCAT
51  CATCGGAGAC GCGCGATGA CGCGGGGCA GCGGTTTGAA GCCTTGAATT
101 GCGCGGCGCA TATGGATGTG GATTGCTGG TCGTCCTCAA CGACAACGAA
151 ATGTCGATTT CCCCCAACGT CGGCGCGTTG CCAAATATC TTGCCAGCAA
201 CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAA GCGCAAACgg
251 GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGagtTTGC CCAAAAAGTC
301 GAACAcaaaa TCAAAACCCT TGCCGAAGAA GCCGAACAG CCAAACAGTC
351 GCTGTCGCTG TTTGAAAATT TCGGCTCCG CTACACC3GC CCCGTGGACG
401 GACACAACGT CGAGAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
451 AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGACCCcg tcaAATACCA CGCCGTGCGc aACCTGCcta
551 AAGAAGGCGG GCGCAAATg ccGTCTGAAA AAGAACCCAA GCCCGCCgCc
601 aaaccgACCT ATACCCAAGT ATTCGGCAAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG

```

```

701 GACTGGTGGG GTTTGAACAA CGATCCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT tacCTTTGCC GCGGGTTTGG CGTGCGAAGG
801 CATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
901 GTCGACCGTG CGGGCATCGT CGGCGCGGAC GGTCCGACCC ATGCCGGCTT
951 GTACGATTTG AGCTTCTTGC GCTGTGTGCC GAACATGATT GTTGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGCTGC TTTGACCTG CTATCAGGCG
1051 GATGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTCAGAC GGCATGAAAA CCGTGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCC TTeatTGCCT TCGGCAGTAT GGTGCGCCAC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTGCGCG ATATGCGCTt
1251 cgtcaaacCG ATAGACGAAG AGTTGATTGT CCGCTTGCC CGAAGCCAcg
1301 accGCATCGT TACCTTGAA GAAAACCGCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTCTTGG AAGTGTGGC GAAACACGGC ATCTGCAAC CCGTTTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGAACA CGGCGATCCG AAAAACTTT
1451 TGGACGATTT GGGTTTGAGT GCCGAAGCGG TGAACGCCG GGTGCGCGAG
1501 TGGCTGCCGG ACCGTGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 033.ng>:

**g033.pep**

```

1 MAAADKLLGG DRRSVAII GD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTAAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDLKDLRSR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKEGGAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLSTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAT
401 ALAVAEKLNA TVADMRVFKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLLGV ADTVTEHGDV KKLLDDLGLS AEAVERRVRE
501 WLPDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

**m033.seq**

```

1 ATGGCGGCGG CAGACAACT CTTGGGCAGC GACCGCCGCA GCGTCGCCAT
51 CATCGCGGAC GCGCGGATGA CGGCGGGGCA GCGGTTTGAA GCCTTGAATT
101 GCGCaG.CGA TATGGATGTr GATTTGCTrG TCGTCCTCAA CGACAACGAA
151 ATGTCGATTT CCCCCAACGT CGGCGCGCTG CCGAAATACC TTGCCAGCAA
201 CGTCGTGCGC GATATGCACG GCCTGTTGAG TACCGTCAAA GCGCAAAACGG
251 GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC
301 GAACACAAAA TCAAAACCCT TGCCGAAGAA GCCGAACAGC CCAACAGTC
351 GCTGTCTTTG TTTGAAAAC TCGGCTTCCG CTACACCGGC CCGTGAGCG
401 GACACAACGT CGAAAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
451 AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGCTACAA
501 ACTCGCCGAA AACGACCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
551 AAGAAAGCGC GCGCGAAATG CCGTCTGAAA AAGAACCCTA GCCGCGCGC
601 AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCATGCGC GAGGGCAGCG
701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT TACCTTTGCC GCGGGTTTGG CTTGCGAAGG
801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTACCCGT TTTGTTTGCC
901 GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTCT
951 GTACGATTTG AGCTTTTTGC GTCGCTGCC GAACATGATT GTCGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGTTGC TTTGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCGTCCG CT...CCGCG GGCACGGGTA CGGGCGCGCC
1101 GGTTCAGAC GGCATGAAAA CCGTGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTGCGCCCC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTGCGCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCTTGCC CGAAGCCAGC

```

214

```

1301 ACCGCATCGT TACCCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTTTGTCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 033>:

```

m033.pep
  1 MAAADKLLGS DRRSVAIIGD GAMTAGQAFE ALNCAXDMDV DLLVVLNDNE
  51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
 101 EHKIITLAE AEHAQSLSL FENFGFRYT PVDGHNVENL VDVLDLRSR
 151 KGPQLLHVIT KKGNGYKLA NDVPKYHAVA NLPKESAAQM PSEKEPKPAA
 201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
 251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA
 301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLSLSTCYQA
 351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGKTA FIAFGSMVAP
 401 ALAVAEKLNA TVADMRVVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
 451 AVLEVLAKHG ICKPVLLGV ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
 501 WLSDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

```

a033.seq
  1 ATGGCGGCGG CGGACAAACA GTTGGGCAGC GACCGCCGCA GCGTCGCCAT
  51 CATCGGCGAC GGCGCGATGA CGGCGGGTCA GCGGTTTGAA GCCTTGAAC
 101 GCGCGGCGCA TATGGATGTG GATTGCTGG TCGTCCTCAA CGACAACGAA
 151 ATGTCGATTT CCCCCAACGT CGGTGCGTTG CCCAAATACC TTGCCAGCAA
 201 CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAA GCGCAAACGG
 251 GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC
 301 GAACATAAAA TCAAACCCCT TCGCGAAGAA GCCGAACACG CCAACAGTC
 351 ACTGTCTTTG TTTGAAAAC TCGGCTTCCG CTATACCGCG CCCGTGGACG
 401 GACACAACGT CGAAAATCTG GTCGATGTAT TGAAGACCT GCGCGGACGC
 451 AAAGGCCCGC AGCTTCTGCA CGTCATCACC AAAAGGGGCA ACGGTACAA
 501 ACTCGCCGAA AACGATCCCG TCAAATACCA CGCGTCGCC AACCTGCCTA
 551 AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCTA GCCGCGCGCC
 601 AAACCGACCT ATACCCAAGT GTTCGGCAA TGGCTGTGCG ACCGGGCGGC
 651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
 701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
 751 ATCGCGGAGC AGCAGCCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG
 801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
 851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGGC
 901 GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGTTT
 951 GTACGATTTA AGCTTTTTCG GCTGCATTCC GAATATGATT GTCGCCGCGC
1001 CGAGCGATGA AAATGAATGC CGCCTGCTGC TTTGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCGTCCG CTATCGCGC GGCACGGGTA CGGGCGTGCC
1101 GGTTCAGAC GGCATGGAAA CCGTGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAACCGCA TTCATTGCCT TCGGCAGTAT GTCGCCCT
1201 GCATTGGCGG TCGCCGAAA ACTGAACGCC ACCGTGCGCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG
1301 ACCGCATCGT TACCCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGCG
1351 GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTCTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 120; ORF 033.a>:

```

a033.pep
  1 MAAADKQLGS DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
  51 MSISPNVGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
 101 EHKIITLAE AEHAQSLSL FENFGFRYT PVDGHNVENL VDVLEDLGR
 151 KGPQLLHVIT KKGNGYKLA NDVPKYHAVA NLPKESAAQM PSEKEPKPAA
 201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
 251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA

```

```

301 VDRAGIVGAD GPTHAGLYDL SFLRCIPNMI VAAPSDENEC RLLSTCYQA
351 DAPAAVRYPR GTGTGVPVSD GMETVEIGKG IIRREGKTA FIAFGSMVAP
401 ALAVAGKLNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGS
451 AVLEVLA KHG ICKPVLLGV ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
501 WLSDRDAAN*

```

m033/a033 98.4% identity over a 509 aa overlap

m033.pep	10	20	30	40	50	60
	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSSISPNVGAL					
a033	MAAADKQLGSDRRSVAIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSSISPNVGAL					
	10	20	30	40	50	60
m033.pep	70	80	90	100	110	120
	PKYLASNVRDMHGLLSTVKAQTKVLDKIPGAMEFAQKVEHKIKTLAEAEHAKQSLSL					
a033	PKYLASNVRDMHGLLSTVKAQTKVLDKIPGAMEFAQKVEHKIKTLAEAEHAKQSLSL					
	70	80	90	100	110	120
m033.pep	130	140	150	160	170	180
	FENFGFRYTGPDGHNVENLVDVLKDLRSRKGPKQLLHVITKKGNKYKLAENDPVKYHAVA					
a033	FENFGFRYTGPDGHNVENLVDVLEDLRGKPKQLLHVITKKGNKYKLAENDPVKYHAVA					
	130	140	150	160	170	180
m033.pep	190	200	210	220	230	240
	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVAITPAMREGSGLVEFEQ					
a033	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVAITPAMREGSGLVEFEQ					
	190	200	210	220	230	240
m033.pep	250	260	270	280	290	300
	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVLF					
a033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVLF					
	250	260	270	280	290	300
m033.pep	310	320	330	340	350	360
	VDRAGIVGADGPTHAGLYDLSFLRCVPNMIVAAPSDENECRLLSTCYQADAPAAVRYPR					
a033	VDRAGIVGADGPTHAGLYDLSFLRCIPNMIVAAPSDENECRLLSTCYQADAPAAVRYPR					
	310	320	330	340	350	360
m033.pep	370	380	390	400	410	420
	GTGTGAPVSDGMETVEIGKGIIRREGKTA FIAFGSMVAPALAVAEKLNATVADMRFVKP					
a033	GTGTGVPVSDGMETVEIGKGIIRREGKTA FIAFGSMVAPALAVAGKLNATVADMRFVKP					
	370	380	390	400	410	420
m033.pep	430	440	450	460	470	480
	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLA KHG ICKPVLLGVADTVTGHGDP					
a033	IDEELIVRLARSHDRIVTLEENAEQGGAGS AVLEVLA KHG ICKPVLLGVADTVTGHGDP					
	430	440	450	460	470	480
m033.pep	490	500	510			
	KKLLDDLGLSAEAVERRVRAWLSDRDAANX					
a033	KKLLDDLGLSAEAVERRVRAWLSDRDAANX					
	490	500	510			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 033 shows 98.4% identity over a 509 aa overlap with a predicted ORF (ORF 033.ng) from *N. gonorrhoeae*:

m033/g033

m033 .pep	MAAADKLLGSDRRSVAIIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSI SPNVGAL	60
	:	
g033	MAAADKLLGSDRRSVAIIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSI SPNVGAL	60
m033 .pep	PKYLASNVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEAEAHAKQSLSL	120
g033	PKYLASNVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEAEAHAKQSLSL	120
m033 .pep	FENFGFRYTGPDVGHNVENLVDVLKDLRSRKGQPLLHVITKKNGYKLAENDPVKYHAVA	180
g033	FENFGFRYTGPDVGHNVENLVDVLKDLRSRKGQPLLHVITKKNGYKLAENDPVKYHAVA	180
m033 .pep	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVAITPAMREGSGLVEFEQ	240
	:	
g033	NLPKEGGAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVAITPAMREGSGLVEFEQ	240
m033 .pep	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA	300
g033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA	300
m033 .pep	VDRAGIVGADGPTHAGLYDLSFLRCVPMIIVAAPSDENECRLLSTCYQADAPAAVRYPR	360
g033	VDRAGIVGADGPTHAGLYDLSFLRCVPMIIVAAPSDENECRLLSTCYQADAPAAVRYPR	360
m033 .pep	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAGFSMVAPALAVAEKLNATVADMRVFKP	420
g033	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAGFSMVATALAVAEKLNATVADMRVFKP	420
m033 .pep	IDIELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTGHGDP	480
g033	IDIELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTEHGDP	480
m033 .pep	KKLLDDLGLSAEAVERRVRAWLSDRDAANX	510
g033	KKLLDDLGLSAEAVERRVREWLPRDAANX	510

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 121>:

g034 .seq

```

1  ATGAGCCGTT  TATGGTTTTT  TGCCGTAAAA  AACATTATAA  TCCGCCTTAT
51  TTACCTATTG  CCCAAGGAGA  CACAAATGGC  ACTCGTATCC  ATGCGCCAAC
101 TGCTTGACCA  CGCCGCCGAA  AACAGCTACG  GCCTGCCCGC  GTTCAACGTC
151 AACAACTCG  AACAAATGCG  CGCCATTATG  GAAGCCGCCG  ACCAAGTCAA
201 CGCGCCCGTC  ATCGTACAGG  CGAGCGCAGG  TCGCGCGCAA  TACGcggGCG
251 CGCCGTTTTT  GCGCCACCTG  ATTCTGGCGG  CAGTCGAAGA  ATTTCCGCAC
301 ATCCCCGTG  TGATGCACCA  AGACCACGGC  GCATCGCCCG  ACGTgtgCCA
351 ACGCTCCATC  CAACTGGGCT  TCTCCTCCGT  GATGATGGAC  GGCTCTTTGC
401 TCGAAGACGG  CAAAACCCCT  TCTTCTTACG  AATACAACGT  CAACGCCACC
451 CGTACCGTCG  TCAACTTCTC  CCACGCCTGC  GGCgtGTCCG  TCGAAGGCGA
501 AATCGGCGTA  TTGGGCAACC  TCGAAACCGG  CGAAGCAGGC  GAAGAAGACG
551 GAGTGGGCGC  GGCAGGCAAA  CTCTCACACG  ACCAAATGCT  CACCAGCGTT
601 GAAGATGCCG  TCGTttTCGT  TAAAGATACC  GCGGTGACG  CATTGGCGAT
651 TGCCGTCGGC  ACCAGCCACG  GCGCATACAA  ATTACCCGT  CCGCCACAG

```



```

701 GCGACGTATT GCGTATCGAC CGCATCAAGG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGgetCCAGC TCCGTTCCGC AAGAatgGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG CAAAGTCAAC
901 ATCGATACCG ACCTGCGCCT CGCTTCCACC GCGCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TTGATCCGCG CAAATACTTG GGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GTTATCTTGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAAT CAAACCTGTT TCGTTGGAAA AAATGGCAAG
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 122; ORF 034.ng>:

**g034.pep**

```

1 MSRLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQSI QLGFSVMMD GSLLEDGKTP SSYEYNVNAT
151 RTVVNFHSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAAGK LSHDQMLTSV
201 EDAVRFVKDT GVDALAIAVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKVI NEYGGNIGET YGVPVEEIVE GIKHGVRKVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL GKTIEAMKQI CLDRYLAFGC
351 EQQAGKIKPV SLEKMASRYA KGELNQIVK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 123>:

**m034.seq (partial)**

```

1 ATGAGCTGTT TATGGTTTTT TGCTGTAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA WACAGCTACG GCyTGCCGGC GTTCAACGTC
151 AACAACTCG WACAGATGCG CGCCATCATG GAGGCTGCAG ACCAAGTCGA
201 CGCCCCCGTC ATCGTACAGG CGAGTGCCGG TCGCGCAAA TATCGGGTG
251 CGCCGTTTTT ACGCCACCTG ATTTTGGCGG CTGTCAAGT ATTTCCACAC
301 ATCCCCGTCG TCATGCACCA AGACCACGGC GCATCACCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TCTCCTCTGT AATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACA
451 CGTACCGTGG TTAACCTTCTC CCACGCTTGC GGCGTATCCG TTGAAGGCGA
501 AATCGGCGTA TTGGGCAACC TCGAAACCGG CGATGCAGGC GAAGAAGACG
551 GTGTAGCGCG AGTGGGCAAA CTTTCCCACG ACCAAATGCT GACCAGCGTC
601 GAAGATGCCG TATGTTTCGT TAAAGATACC GCGTTGACG CATTGGCTAT
651 TGCCGTCGGC ACCAGCCACG GCGCATACAA ATTCACCCGT CCGCCCACAG
701 GCGATGTATT ACGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA C...

```

This corresponds to the amino acid sequence <SEQ ID 124; ORF 034>:

**m034.pep (partial)**

```

1 MSCLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE XSYGLPAFNV
51 NNLXQMRAIM EAADQVDAPV IVQASAGARK YAGAPFLRHL ILAAVEVFPH
101 IPVVMHQDHG ASPDVCQSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFHSHAC GVSVEGEIGV LGNLETGDAG EEDGVGAVGK LSHDQMLTSV
201 EDAVCFVKDT GVDALAIAVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMH...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 125>:

**a034.seq**

```

1 ATGAGCCGTT TATGGTTTTT TGCCGCAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA AACAGCTACG GCCTGCCCCG GTTCAACGTC
151 AACAACTCG AACAAATGCG CG.ATTATG GAAGCCGCGC ACCAAGTCAA
201 CGCGCCCGTC ATCGTACAGG CGAGCGCAGG TCGCGCAAA TACGCGGGCG
251 CGCCGTTTTT GCGCCACCTG ATTTTGGCGG CTGTCAAGA ATTTCCGCAC
301 ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCC ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TTTCTCTCGT GATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTATG AATACAACGT CAACGCCACG

```

```

451 CGTACCGTGG TTAATTTCTC CCACGCCTGC GCGGTATCCG TTGAAGGCGA
501 AATCGGCGTA TTGGGCAACC TCGAAACTGG CGAAGCCGGC GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCCACG ACCAAATGCT CACCAGCGTC
601 GAAGATGCCG TCGGTTTCGT TAAAGATACC GCGGTTGACG CATTTGGCGAT
651 TGCCGTCGCG ACCAGCCACG GCGCGTACAA ATTCAACCGT CCGCCCACAG
701 GCGACGTGTT GCGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CCGCTCCAGC TCCGTTCCGC AAGAATGGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CCGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG TAAAGTCAAC
901 ATCGATACCG ACTTGGCGCT TGCTTCCACC GCGCGGTAC GCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TCGATCCGCG CAAATATTG AGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCTTCGACC GCTACCTCGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAAT CAAACCGGTT TCCTTGGAAG AAATGGCAAA
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 126; ORF 034.a>:

```

a034.pep
  1 MSRLWFFAAK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
 51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCORSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFHSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAVGK LSHDQMLTSV
201 EDAVRFVKDT GVDALAIAGV TSHGAYKFTR PPTGDLVRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKV NEYGGNIGET YGVPVEEIVE GIKHGVKRVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL SKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMANRYA KGELNQIVK*

```

m034/a034 96.9% identity over a 257 aa overlap

```

              10      20      30      40      50      60
m034.pep    MSCLWFFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM
              |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a034        MSRLWFFFAAKNIIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM
              10      20      30      40      50      60

              70      80      90     100     110     120
m034.pep    EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCORSI
              |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a034        EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCORSI
              70      80      90     100     110     120

              130     140     150     160     170     180
m034.pep    QLGFSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFHSHACGVSVEGEIGVLGNLETGDAG
              |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a034        QLGFSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFHSHACGVSVEGEIGVLGNLETGEAG
              130     140     150     160     170     180

              190     200     210     220     230     240
m034.pep    EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAGVTSHGAYKFTRPPTGDLVRID
              |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a034        EEDGVGAVGKLSHDQMLTSVEDAVRFVKDTGVDALAIAGVTSHGAYKFTRPPTGDLVRID
              190     200     210     220     230     240

              250
m034.pep    RIKEIHQALPNTHIVMH
              |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a034        RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVKRVN
              250     260     270     280     290     300

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 034 shows 96.5% identity over a 257 aa overlap with a predicted ORF (ORF 034.ng) from *N. gonorrhoeae*:

## m034/g034

m034 . pep	MSCLWFFAVKNI IIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM	60
g034	MSRLWFFAVKNI IIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNMLEQMRAIM	60
m034 . pep	EAADQVDAPVIVQASAGARKYAGAPFLRHILAAVEVFPHIPVVMHQDHGASPDVCQRSI	120
g034	EAADQVNAPVIVQASAGARKYAGAPFLRHILAAVEEFPHIPVVMHQDHGASPDVCQRSI	120
m034 . pep	QLGFSSVMDGSLMEDGKTPSSYEYNVNATRTVVNFSSHACGVSVGEIGVLGNLETGDAG	180
g034	QLGFSSVMDGSLLEDGKTPSSYEYNVNATRTVVNFSSHACGVSVGEIGVLGNLETGEAG	180
m034 . pep	EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAGVTSHGAYKFRPPTGDVLRID	240
g034	EEDGVGAAGKLSHDQMLTSVEDAVRFVKDTGVDALAIAGVTSHGAYKFRPPTGDVLRID	240
m034 . pep	RIKEIHQALPNTHIVMH	257
g034	RIKEIHQALPNTHIVMHGSSSVPOEWLKVINEYGGNIGETYGVPVEEIVEGIXHGVKVN	300

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 127>:

## g036 . seq

1	ATGCTGAAGC	CGTGTTCGTT	ATACAGTGCC	TGTGCGGCGG	cgttgcCTGC
51	GCGGACTTCG	AGCAGCAGGC	GTTGCGTGCC	TTCGGGCAGA	TGTGCGTACC
101	AATATTCGAG	CAGGGCGGAC	GCAACGCCCC	GTCGGCGGCA	TTCGGGCGCG
151	GTGGCAATCA	GGTGCAATTC	GGATTTCGTC	GGCAGGTTCT	GCCAAACGAT
201	AAAGCGGCA	ATCCTGCCGT	CTTTTCCGC	AAGGAAAACC	TGTTCGGACG
251	GCGAAACAAG	CGCGGACTCA	AATTGGCGTT	GCGTCCACGC	GGACGGGTTG
301	CAGACGGTAT	CGAGCGCGGC	CAGTGCGGCG	CAGTCGGACG	GTGAGGCTGG
351	GCGGATGTTT	ATGTTTCGTG	CTTCCGTTCC	GCCTGTTCTT	TGGCAGTCAG
401	GGCGATTTTG	TTGCGGACGT	AGAGCAGTTC	GGCGTGTGCC	GCGCCAGTTG
451	CGGGATAGCC	GCCGCCGAGG	GCGAGCGCGA	GAAAATCGGC	GGCGGTCCGC
501	ATATCGGGTT	TGCCTGAGAA	GGGCGGACGG	TTTTCAGTGC	CGAACGCAC
551	GCCGATGCCG	TCTGAAAAGA	CGTACCCCTC	GGGGAGGGCA	ATGTCTGCCG
601	CCCTACCGAC	TTGATAATCG	CTCAAACGGC	GGCGGTTCAG	CGTGTCCGAC
651	CACGCATAAA	ACACTTCGCC	CATACGCGCG	TCCGCAGCGG	CGAGTATGCA
701	GCTTTGCGGC	GGCGGCAGCG	AGGCGCGCGC	ATCGAGCGTG	GGGATGCCGA
751	TTAAAGGCGT	GTCGAACGGC	GTTGCCAAAC	CTTGCGCCAC	GCCGATGCCG
801	ATACGCAGTC	CGGTAA			

This corresponds to the amino acid sequence <SEQ ID 128; ORF 036.ng>:

## g036 . pep

1	MLKPCLVYSA	CAALPARTS	SSRRCVPSGR	CAYQYSSRAD	ATPRRHSGA
51	VAIRCSSDSS	GRFCQTIKAA	ILPSFSARKT	CSDGETSADS	NWRCVHADGL
101	QTVSSAASAA	QSDGEAGRMF	MFVPSVPPVL	WQSGRFCCGR	RAVRRVPRQL
151	RDSRRRGRAR	ENRRRSAYRV	CLRRADGFPV	RTHCRCLKR	RTPRGGQCLP
201	PYRLDNRSNG	GGACRTTHK	TLRPYARPQR	RVCSEAAAAA	RRRHRAWGCR
251	LKACRTALPN	LAPRRCRYAV	R*		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 129>:

## m036 . seq

1	ATGCTGAAGC	CGTGCGCCGT	GTACAGTGCC	TGTGCGGCGG	TGTTGCCTGC
51	ACGGACTTCG	AGCAGCAGGC	GTTGCGTGTC	TTCGGGCAGA	TGTGTGAACC
101	AATATTCGAG	CAGGGCGGAC	GCAATTCCTT	GCGGCGGCA	TTCGGGCGCG
151	GTGGCAATCA	GGTGCAATTC	GGATTTCGTC	GGCAGGTTCT	GCCAAACGAT
201	AAAGCGGCA	ATCCGg.CGT	CTTTTCCGC	AAGGAAAACC	TGTTCGGACG
251	GCGAAACCAG	TGCGGACTCA	AATTGGCGTT	GCGTCCATGC	GGACGGGTTG
301	CAGACGGCAT	CGAGTGCGGC	CAGCTCCTCA	CAATCGGCAC	AAACGGCAGC

220

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTCT GCCTSTTCTT TGGCAGTCAG
401 GGGGATTTTG TTGCGGACGT AGAGCAAACC GCGGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA ACACTTCGCC CATAAGAGCG TCCGTAGCGG CAAGGATGCA
701 GCTTTGCGGC GGCAGGACGG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGGGGT ATCAAACGGC GTTGCCAAAC CCTGAGCTAC ACCGATGCCG
801 ATACGCAGTC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 130; ORF 036>:

```

m036.pep
1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPXSFSAKKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLRLKR RIPPAAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*R QGCSFAAAAA RRRHRARVRR
251 LRGYQTALPN PELHRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 131>:

```

a036.seq
1 ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTCCTGCG
51 ACGGACTTCG AGCAGCAGGC GTTGCGTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCCTT GCGGCGCGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCTGCG GGCAGGTCTT GCCAAACGAT
201 AAAGGCGGCA ATCCCGCCGT CTTTTTCCGC AAGGAAAACC TGTTCGACG
251 GCGAAACCAG TCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGCAT CGAGCGCGGC GAGTGCGGCG CAATCGGCAT AAACGGCGCG
351 GCGGATGTTT ACAGGCGCGC CCTCCGTTCC GCCTGTCTTT TGGCAGTCAA
401 GCGGATTTTG TTGCGGACGT AGAGCAGCTC G3CGTGTGCC GCAGCGACGG
451 CCGGAAAACC GCCTTCAGCC GCCAGATTGA GGAAGTCGGC GCGGTCGGC
501 ATATCGGGTT TGCCTGAGAA GGCAGGACGG TTTTCCAGCG CGAACGCATT
551 GCCGATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGACCGAC CTGATAATCG CTCAAACGGC GCGGTTTCAG CGTGTGGAAC
651 CATGCATAAA ACACTTCGCC CATAAGTGGC TCCGAGCGCG CAAGGATGCA
701 GCTTTGCGGC GGCAGGACGG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGAGT ATCAAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801 ATACGCAGTC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 132; ORF 036.a>:

```

a036.pep
1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPPSFSAKKT CSDGETSADS NWRCVHADGL
101 QTASSAASAA QSA*TARRMF TGAPSVPPVL WQSRRFCCGR RAARRVPQRR
151 RENRLQPPD* GSRRRSAYRV CLRRADGFFA RTHCRCLKR RILPAAGCLP
201 PDRPDNRSNG GGSACRTMHK TLRPYVRPQR QGCSFAAAAA RRRHRARVRR
251 LKEYQTALPN LAPRRCRYAV P*

```

m036/a036 85.6% identity over a 270 aa overlap

```

10 20 30 40 50 60
m036.pep MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIIPWRRHSGAVAIIRCSSDSS
|||||
a036 MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAIIPWRRHSGAVAIIRCSSDSS
10 20 30 40 50 60

70 80 90 100 110 120
m036.pep GRFCQTIKAAIPXSFSAKKTCS DGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF
|||||
a036 GRFCQTIKAAIPPSFSAKKTCS DGETSADSNWRCVHADGLQTASSAASAAQSAXTARRMF
70 80 90 100 110 120

```

221

	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA					
	:   :      :        :					
a036	TGAPSVPPVLWQSGRRFCCGRRRAARRVPQRRRENRLQPPDXGSRRSAYRVCLRRADGFPA					
	130	140	150	160	170	180
m036.pep	190	200	210	220	230	240
	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPXRGCSFAAAAA					
	:					
a036	RTHCRCLKRRILPAAGCLPPDRPDNRSNGGGSACRTMHKTLRPYVRPQRQCSFAAAAA					
	190	200	210	220	230	240
m036.pep	250	260	270			
	RRRHRARVRLRGYQTALPNPELHRCRYAVRX					
	:        :					
a036	RRRHRARVRLKEYQTALPNLAPRRCRYAVPX					
	250	260	270			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 036 shows 74.9% identity over a 271 aa overlap with a predicted ORF (ORF 036.ng) from *N. gonorrhoeae*:

m036/g036

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAI PWRRHSGAVAIRCSSDSS					
g036	MLKPCLVYSACAAALPARTSSSRRCVPSGRCAQYSSRADATPRRRHSGAVAIRCSSDSS					
	10	20	30	40	50	60
m036.pep	70	80	90	100	110	120
	GRFCQTIKAAIPXSFSARKTCS DGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF					
g036	GRFCQTIKAAILPSFSARKTCS DGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF					
	70	80	90	100	110	120
m036.pep	130	140	150	160	170	180
	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA					
	:                    :   :    :            :					
g036	MEVPSVPPVLWQSGRFCCGRRAVRRVPRLRDSRRRGRARENRRRSAYRVCLRRADGFVP					
	130	140	150	160	170	180
m036.pep	190	200	210	220	230	240
	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPXRGCSFAAAAA					
	:          :					
g036	RTHCRCLKRRTPRGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPQRRVCSFAAAAA					
	190	200	210	220	230	240
m036.pep	250	260	270			
	RRRHRARVRLRGYQTALPNPELHRCRYAVRX					
	:        :					
g036	RRRHRAWGCRKACRTALPNLAPRRCRYAVRX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 133>:

m036-1.seq

```

1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTTGCCTGC
51  ACGGACTTCG AGCAGCAGGC GTTCTGTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCGCGCCT CTTTTCCTCG AAGGAAAACC TGTTGCGGACG
251 GCGAAACCAG TCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301 CAGACGGCAT CGAGTGC GAGCTCCTCA CAATCGGCAC AAACGGCAGC

```

222

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTCT GCCTGTCTTT TGGCAGTCAG
401 GGCGATTTTG TTGCGGACGT AGAGCAAACC GCGGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCTCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA ACACTTCGCC CATAAGAGCG TCCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 134; ORF 0036-1>:

m036-1.pep

```

1 MLKPCAVYSA CAAVLPART SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDCETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPLWPMR ESRRQSAYPV CLRTAELLPA RTRCLRLKR RIPPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*

```

m036-1/g036 76.8% identity in 228 aa overlap

	10	20	30	40	50	60
m036-1.pep	MLKPCAVYSACAAVL	PARTSSSRRCVSSGR	CVNQYSSRADAI	PWRRHSGAVAI	RCSSDSS	
g036	MLKPCLVYSACAAAL	PARTSSSRRCVPSGR	CAYQYSSRADAT	PRRRHSGAVAI	RCSSDSS	
	10	20	30	40	50	60
m036-1.pep	GRFCQTIKAAIPPS	FARKTCS	DGETSADSNWRC	VHADGLQTASSA	ASSSQSAQTAR	MF
g036	GRFCQTIKAAILPS	FARKTCS	DGETSADSNWRC	VHADGLQTVSSA	ASAAQSDGEAG	RMF
	70	80	90	100	110	120
m036-1.pep	TGALSVRPVLWQSG	RFCCGRANRRVR	HGRQDNRPLWPM	RESRRQSAYPV	CLRTAELLPA	
g036	MFVPSVPPVLWQSG	RFCCGRVRRVPR	QLRDSRRRGRAR	ENRRRSAYRV	CLRRADGFPV	
	130	140	150	160	170	180
m036-1.pep	RTRCLRLKRRIPPA	AGCLPPARPDNR	SNGGSSAYRTMH	KTLRPYERPX		
g036	RTHCRCLKRRTPR	GGQCLPPYRLDN	RNSNGGSACRTH	KTLRPYARPQR	RVCSFAAAAA	
	190	200	210	220	229	
m036-1.pep						
g036						
	190	200	210	220	230	240
m036-1.pep						
g036						
	250	260	270			
m036-1.pep						
g036						

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 135>:

g038.seq

```

1 ATGACTGATT TCCGCCAAGA TTCTCTCAAA TTCTCCCTCG CCCAAAATGT
51 TTTGAAATTC GGCGAATTTA CCACCAAAGC CGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTC AACGACGGCG CGTCCACGCT GCAACTGGCA
151 AAATTCTATG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC GCGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCAATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGCG CGGTGTCGCC
451 ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTAATTTGT CCGCGTTTCA
501 GGAAGTGGA AAACAATACG GCCTGCCCGT CGCCCCATC GCCAGCCTGA
551 ACGATTGTT TATCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
601 GAACCCGTCC GCACCTACCG CCGGCAGTAC GCGGTAGAAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 136; ORF 038.ng>:

```
g038.pep
1  MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGASTLQLA
51  KFYAQSIIES GIRFDMLEFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101 AKDRGEGGVL VGAPLKGRVL IIDVISAGT SVRESIKLIE AEGATPAGVA
151 IALDRMEKGT GKLSAVQEVE KQYGLPVAPI ASLNDFILL QNNPEFGQFL
201 EPVRTYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 137>:

```
m038.seq
1  ATGACCGATT TCCGCCAAGA TTCTCCTCAA TTCTCCCTCG CCCAAAATGT
51  TTTGAAATTC GCGGAATTTA CCACCAAGGC AGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
151 AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGTCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC ACGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGC CGGTGTCGCC
451 ATCGCGCTCG ATCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
501 GGAAGTGGA AAACAATACG GkCTGCCCGT CGCCCCATC GCCAGCCTGA
551 ACGATTGTGT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
601 GAACCCGTCC GAGCCTACCG TCGGCAGTAC GCGGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 138; ORF 038>:

```
m038.pep
1  MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
51  KFYAQSIIES GIRFDMLEFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101 AKDHGEGGVL VGAPLKGRVL IIDVISAGT SVRESIKLIE AEGATPAGVA
151 IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDFILL QNNPEFGQFL
201 EPVRAYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 139>:

```
a038.seq
1  ATGACCGATT TCCGCCAAGA TTCTCCTCAA TTCTCCCTCG CCCAAAATGT
51  TTTGAAATTC GCGGAATTTA CCACCAAGGC CGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
151 AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC ACGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGC CGGTGTCGCC
451 ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
501 GGAAGTGGA AAACAATACG GCCTGCCCGT CGCCCCATC GCCAGCCTGA
551 ACGATTGTGT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
601 GAACCCGTCC GAGCCTACCG TCGGCAGTAC GCGGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 140; ORF 038.a>:

```
a038.pep
1  MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
51  KFYAQSIIES GIRFDMLEFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101 AKDHGEGGVL VGAPLKGRVL IIDVISAGT SVRESIKLIE AEGATPAGVA
151 IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDFILL QNNPEFGQFL
201 EPVRAYRRQY GVE*
```

m038/a038 100.0% identity over a 213 aa overlap

```

          10      20      30      40      50      60
m038.pep  MTDFRQDFLK FSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
```

224

```

|||||
a038      MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
           10      20      30      40      50      60

           70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
           |||||||
a038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
           70      80      90      100     110     120

           130     140     150     160     170     180
m038.pep  IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELS AVQVEVEKQYGLPVAPI
           |||||||
a038      IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELS AVQVEVEKQYGLPVAPI
           130     140     150     160     170     180

           190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
           |||||||
a038      ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
           190     200     210

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 038 shows 98.1% identity over a 213 aa overlap with a predicted ORF (ORF 038.ng) from *N. gonorrhoeae*:

**m038/g038**

```

           10      20      30      40      50      60
m038.pep  MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
           |||||||
g038      MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGASTLQLAKFYAQSIIES
           10      20      30      40      50      60

           70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
           |||||||
g038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDRGEGGVLVGAPLKGRVL
           70      80      90      100     110     120

           130     140     150     160     170     180
m038.pep  IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELS AVQVEVEKQYGLPVAPI
           |||||||
g038      IIDDVISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGKLS AVQVEVEKQYGLPVAPI
           130     140     150     160     170     180

           190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
           |||||||
g038      ASLNDLFILLQNNPEFGQFLEPVRTYRRQYGVEX
           190     200     210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 141>:

**g039.seq**

```

1  ATGCCGTC CG AACCACCTGC CGCTTCAGAC GGCATCAAAC CGACACACAC
51  CGAGAAAACA TCATGCCCGC CTGTTTCTGT CCGCACTGCA AAACCCGCCT
101 CTGGGTCAAA GAAAcccagC TCAAcgtCgC ccaagGCTTC GTCGTCTgcc
151 aaAAAtgcga agGGCTgttt aaAgccaaaG accAtctggc aaGcacGAAA
201 gaacctatat tcaacgattg gcccgaaagt gtttcgggat gTcaaaCTCG

```



225

```

251 TCcaccgcaT cggcaccgcac gccattagca aGAaacagat gtccccgcgac
301 gaaatCgccg atatacctcaa cggcgggtaca acCCTGCACG ATACGCCGCC
351 CGCAACCGCC GCTGCCGCac ctGCCGCCGC ACCGcaggTT TCCGTACCGC
401 CCGCCCGTCA GGAAGGGCTC AACTGGACTA TTGCAACCCT GTTCGCACTT
451 ATCGTCCTCA TTATGCAGCT TTCCTACCTC TTCATCCTAT GA

```

This corresponds to the amino acid sequence <SEQ ID 142; ORF 039.ng>:

g039.pep

```

1 MPSEPPAASD GIKPTHTEKT SCPPVSVRTA KPASGSKKPS STSPKASSSA
51 KNAKGCLPKP TIWQARKNLY STIGPKLFRD VKLVHRIGTH AISKKQMSRD
101 EIADILNGGT TLHTTTPATA AAAPAAAPQV SVFPAQEGE NWTIATLFAL
151 IVLIMQLSYL FIL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 143>:

m039.seq

```

1 ATGCCGTCCG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51 CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCCTGCA AAACCCGTCT
101 CTGGGTCAAA GAAACCAAC TCAATGTCGC CGnnnnnnnnn nnnnnnnnnn
151 nnnnnnnnnn nnnr nnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
201 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnCCC GAGGCTGTTT
251 CGGATGTCAA ACTCGTTCAC CGTATCGGCA CGCGCGCCAT CGGCAAGAAA
301 CAGATTTCCT GTGACGAAAT CGCCGGCATC CTCAACGGCG GTACAACCCA
351 GCCCGATATT CCGCCCGCAA CCGCCGCCAC CCCTGCTGCC GCACCGCAGG
401 TTACCGTACC GCCCGCCGCG CCGCCCGGTC AGGATGGGTT CAACTGGACG
451 ATTGCAACCC TGTTTGCCCT TATCGTCCTC ATTATGCAGC TTTCTACCT
501 CGTCATCCTA TGA

```

This corresponds to the amino acid sequence <SEQ ID 144; ORF 039>:

m039.pep

```

1 MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPPXXXXX
51 XXXXXXXXXX XXXXXXXXXX XXXXXXXXXP EAVSDVKLVH RIGTRAIGKK
101 QISRDEIAGI LNGGTTQPD I PPATAATPAA APQVTVPPAA PARQDGFNWT
151 IATLFALIVL IMQLSYLVIL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 145>:

a039.seq

```

1 ATGCCGTCTG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51 CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCCTGCA AAACCCGTCT
101 CTGGGTCAAA GAAACCAAC TCAATGTCGC CCAAGGCTTC GTCGTCTGCC
151 AAAAATGCGA AGGAATGTTT AAAGCCAAAG ACCATCTGGC AAGCAGAAA
201 GAACCCATAT TCAACGATT. TGCCCGAAGC TGTTTCGGAT GTCAAACCTCG
251 TTCACCGCAT CGGCACGAGC GCCATCGGCA AGAAACAGAT TTCCCGTGAC
301 GAAATCGCCG GCATCCTCAA CGGCGGCACA ACCCAGCCCG ATATTCCGCC
351 CGCAACCGCC GCCACCCCTG CTGCCGCACC GCAGGTTACC GTACCGCCCG
401 CCGCGCCCGC CCGTCAGGAT GGGTTCAACT GGACGATTGC AACCCGTGTT
451 GCCCTTATCG TCCTCATTAT GCAGCTTTCC TACCTCGTCA TCCTATGA

```

This corresponds to the amino acid sequence <SEQ ID 146; ORF 039.a>:

a039.pep

```

1 MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPKASSSA
51 KNAKECLPKP TIWQARKNPY STIXPEAVSD VKLVHRIGTS AIGKKQISRD
101 EIAGILNGGT TQPDIPPATA ATPAAAPQVT VPPAAPARQD GFNWTIATLF
151 ALIVLIMQLS YLVIL*

```

m039/a039 79.4% identity over a 170 aa overlap

```

              10      20      30      40      50      60
m039.pep      MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNMSMPXXXXXXXXXXXXXXXXX
              |||||
a039           MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNMSMPKASSSAKNAKECLPKP

```

226

	10	20	30	40	50	60
m039.pep	70	80	90	100	110	120
	XXXXXXXXXXXXXXXXXXXX	PEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPD				
a039	70	80	90	100	110	
	TIWQARKNPYSTIK----	PEAVSDVKLVHRIGTSAIGKKQISRDEIAGILNGGTTQPD				

	130	140	150	160	170
m039.pep	PPATAATPAAAPQVTVPPA	PARQDGFNWTIATLFALIVLIMQLSYLVILX			
a039	PPATAATPAAAPQVTVPPA	PARQDGFNWTIATLFALIVLIMQLSYLVILX			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 039 shows 60.8% identity over a 171 aa overlap with a predicted ORF (ORF 039.ng) from *N. gonorrhoeae*:

m039/g039

	10	20	30	40	50	60
m039.pep	MPSEPPYASDGIKPDTHEE	ICPPVVSAPTAKPVSGSKKPN	SMSPXXXXXXXXXXXXXXXX			
g039	MPSEPPAASDGIKPTHTEK	TSCPPVSVRTAKPASGSKKPS	SSTSFKASSSAKNAKGLKPK			

	10	20	30	40	50	60
m039.pep	70	80	90	100	110	120
	XXXXXXXXXXXXXXXXXXXX	PEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPD				
g039	70	80	90	100	110	
	TIWQARKNLYSTIG----	PKLFRDVKLVHRIGTHAISKQMSRDEIADILNGGTTLHDT				

	130	140	150	160	170
m039.pep	PPATAAT-PAAAPQVTVPPA	PARQDGFNWTIATLFALIVLIMQLSYLVILX			
g039	PPATAAAAPAAAPQVSVPPA	---RQEGLNWTIATLFALIVLIMQLSYLFILX			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 147>:

g040.seq

```

1  ATGAACGCGC  CCGACAGCTT  TGTCGCCAC  TTCCGCGAAG  CCGCCCCCTA
51  CATCCGCCAA  ATGCGCGGCA  CGACACTGGT  CGCCGGCATA  GAcggCCGCC
101 TGCTCGAAGG  CGGCACCTTA  AATAAGCTCG  CCGCCGACAT  CGGGCTGTTG
151 TCGCAACTGG  GCATCCGACT  CGTCCTCATC  CACGGCGCGT  ACCACTTCCT
201 CGAcgCCTC  GCCGCCGCGC  AAGgccGCAC  GCCGCATTAT  TGCCGgggtt
251 tGCGCGTTAC  CGACGaAACc  tcGctcgGAC  AGGCGCAGCA  GtttGCCGGC
301 AccgTCCGCA  GCCGTTTTGA  agcCGCATTG  tgcggcagCG  tttcaggatt
351 cgcgCGCGCG  CCTTCCGTCC  CGCTCGTAtc  gggcaacttc  ctgacCGCCC
401 GTCcgatggg  cgtgattgac  ggaACCGata  tggaaatacgc  gggggttattc
451 cgcaaaaccg  ACACCGCCGC  CCTCCGTTTC  CAACTCGACG  CGGGCAATAT
501 CGTCTGGATG  CCGCCGCTCG  GGCATTCTTA  CGGCGGCAAA  ACCTTCAATC
551 TCGATATGGT  GCAGGCCGCC  GCTTCCGTCT  CCGTCTCGCT  TCAGGCCGAA
601 AAACCTCGTT  ACCTGACCTT  TTCAGACGGC  ATTTCCCGCC  CCGACGGCAC
651 GCTCGCCGAA  ACCCTCTCGG  CACAGGAAGC  GCAATCGCTG  GCGGAACACG
701 CCGCCAGCGA  AACCAGACGA  CTGATTTCGT  CCGCCGTGTC  CGCGCTCGAA
751 GGCGGCGTGC  ATCGCGTCCA  AATCCTCAAC  GGGGCCGCCG  ACGGACGCTT
801 GCTGCAAGAA  CTCTTACCCC  GCAACGGCAT  CGGCACGTCC  ATTGCCAAAG
851 AAGCCTTCGT  CTCCATCCGG  CAGGCGCACA  GCGGCACAT  CCCGCACATC
901 GCCGCCCTCA  TCCGCCCGCT  GGAAGAACAG  GGCGTCCTAT  TGCACCGCAG
951 CCGCGAATAC  CTCGAAAACC  ACATTTCCGA  ATTTTCCATC  CTCGAACACG

```

```

1001  ACGGCGACCT GTACGGCTGT GCCGCACTCA AAACCTTTGC CGAAGCCGAT
1051  TCGGCGGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGG
1101  ctACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151  GCATAAGCAG GCTGTTCGCA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201  GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGCTGCCCG AAACGCGGCG
1251  CAAAGACTAC CGCAGCAACG GACGAAACCC GCATATTCTG GTGCGTCGCC
1301  TGCACCGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 148; ORF 040.ng>:

```

g040.pep
1  MNAPDSFVAH FREAAPYIRQ MRGTTLVAGI DDRLLEGGTL NKLAADIGLL
51  SQLGIRLVLI HGAYHFLDRL AAAQGRTPHY CRGLRVTDET SLGQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPMGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWM PPLGHSYGGK TFNLDMVQAA ASVAVSLQAE
201 KLVYLTLSDG ISRPDGTAE TLSAQEAQSL AEHAASETRR LISSAVAAL
251 GGVHRVQILN GAADGSLQSE LFRNGIGTS IAKEAFVSIR QAHSQDIPHI
301 AALIRPLEEQ GVLLHRSREY LENHISEFSI LEHDGDLYGC AALKTFEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRRKDY RSNRNPIL VRRLHR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 149>:

```

m040.seq
1  ATGAGCGGCG CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGTCCCCTA
51  CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
101 TGTCGGAAGG TGATACCTTA AACAAGCTCG CCGCCGACAT CGGGCTGTTG
151 TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
201 CGACCGCCAC GCCGCGCTC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
251 TGC CGCTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTCGCGGC
301 ACCGTCCGCA GCCGTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
351 CGCGCGCGCG CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
401 GTCCGATAAG TGTGATTGAC GGAACCGATA TGGAATACGC GGGCGTTATC
451 CGCAAAACCG ACACCGCGCG CCTCCGTTTC CAACTCGACG CGGGCAATAT
501 CGTCTGGCTG CCGCCGCTCG GACATTCTTA CAGCGGCAAG ACCTTCTATC
551 TCGATATGCT TCAAACCGCC GCCTCCGCGC CGTCTCGCT TCAGGCGCAA
601 AAACCTCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACCGCAC
651 GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
701 CCGCGGGGCA AACCGCACGG CTGATTTCGT CCGCCGAACT CTTCACCCGC
751 AACGGCATCG GCACGTCCAT TGCCAAAGAA GCCTTCGTCT CCATCCGGCA
801 rGCGCAywgG G.CGACATCC CGCACATCGC CGCCCTCATC CGCCCGCTGG
851 AAGAACAGGG CATCCTGCTG CACCGCAs.c GCGAATACCT CGAAAACAC
901 ATTTCCGAAT TTTCCATCCT CGAACACGAC GGCAACCTGT ACGGTTGCGC
951 CGCCCTGAAA ACCTTTGCCG AAGCCGATTG CGGCGAAATC GCCTGCCTTG
1001 CCGTCTCGCC GCag.cACAG GACGGCGGCT ACGGCGAACG CnTGCTTGCC
1051 CACATTATCG ATAAGGCGCG CGGCATAGGC ATAAGCAGGC TGTTGCGACT
1101 GTCCACAAAT ACCGGCGAAT GGTTTGCCGA ACGCGGCTTT CAGACGGCAT
1151 CGGAAGACGA GTTGCCCGAA ACGGCGGCGA AAGACTACCG CAGCAACGGA
1201 CGGAACCTGC ATATTCTGGT ACGTCGCCTG CACCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 150; ORF 040>:

```

m040.pep
1  MSAPDLFVAH FREAVPYIRQ MRGKTLVAGI DDRLLEGDTL NKLAADIGLL
51  SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFYLDMLQTA ASAAVSLQAE
201 KLVYLTLSDG ISRPDGTAE TLSAQEAQSL AEHAGGQTRR LISSAELFTR
251 NGIGTSIAKE AFVSIRQAHX XDIPHIAALI RPLEEQGILL HRXREYLENH
301 ISEFSILEHD GNLYGCAALK TFAEADCGEI ACLAVSPQXQ DGGYGERXLA
351 HIIDKARGIG ISRLFALSTN TGEWFAERGF QTASEDELPE TRRKDYRSNG
401 RNSHILVRRL HR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 151>:

a040.seq

```

1  ATGATCGTGC CCGACCTCTT TGTGCCCCAC TTCCGCGAAG CCGCCCCCTA
51  CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCCGCATA GACGACCGCC
101 TGCTCGAAGG TGATACCTTA AACAAATTTCG CCGCCGACAT CGGGCTTTTG
151 TCGCAACTGG GCATCAGGCT CGTCTCATC CACGGCGCGC GCCACTTCCT
201 CGACCGCCAC GCCGCGCGC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
251 TGC CGCTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTGCCGGC
301 ACCGTCCGCA GCCGTTTTGA AGCCGCATTG TGCGGCAGCG TTTCGGGTT
351 CGCGCGCGCG CCTTCCGTCC CGTCTGTATC GGGCAACTTC CTGACCGCCC
401 GTCCGATAGG TGTGATTGAC GGAACCGATA TGGAAATACGC GGGCGTTATC
451 CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGCAATAT
501 CGTCTGGCTG CCGCCGCTCG GACATTCTTA CAGCGGCAAG ACCTTCCATC
551 TCGATATGCT TCAAACCGCC GCCTCCGTCG CCGTCTCGCT TCAGGCCGAA
601 AACTCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
651 GCTCGCCGTA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
701 CCGGCGGCGA AACGCGACGG CTGATTTCGT CCGCCGTTGC CGCGCTCGAA
751 GGCGGCGTGC ATCGCGTCCA AATCCTCAAC GGAGCCGCCG ACGGCAGCCT
801 GCTGCAAGAA CTCTTCACCC GCAACGGCAT CGGCACGTCC ATTGCCAAAG
851 AAGCCTTCGT CTCCATCCGG CAGGCGCACA GCGGCGACAT CCCGCACATT
901 GCCGCCCTCA TCCGCCCGCT GGAAGAACAG GGCATCCTGC TGCACCGCAG
951 CCGCGAATAC CTCGAAAACC ACATTTCCGA ATTTTCCATC CTCGAACACG
1001 ACGGCAACCT GTACGGTTGC GCCGCCCTGA AAACCTTTGC CGAAGCCGAT
1051 TGCGGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGG
1101 CTACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151 GCATAAGCAG GCTGTTTGCA CTGTCCACAA ATACCGGCGA ATGTTTGCC
1201 GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGTTGCCCG AAACGCGGCG
1251 CAAAGACTAC CGCAGCAACG GACGGAACTC GCATATTCTG GTGCGTCGCG
1301 TGCACCGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 152; ORF 040.a>:

```

a040.pep
1  MIVPDLFVAH FREAAPYIRQ MRGKTLVAGI DDRLLEGDTL NKFAADIGLL
51  SOLGIRLVLI HGARHFIDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFHLDMLQTA ASVAVSLQAE
201 KLVYLTLSDG ISRPDGT LAV TLAQEAQSL AEHAGGETRR LISSAVAAL
251 GGVHRVQILN GAADGSLLE LFRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GILLHRSREY LENHISEFSI LEHDGNLYGC AALKTFEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRRKDY RSNGRNSHIL VRRLRH*

```

m040/a040 91.5% identity in 436 aa overlap

```

10      20      30      40      50      60
m040.pep MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTL NKLAADIGLLSOLGIRLVLI
| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
a040      MIVPDLFVAHFREAAPYIRQMRGKTLVAGIDDRLLEGDTL NKFAADIGLLSOLGIRLVLI
10      20      30      40      50      60

70      80      90      100     110     120
m040.pep HGARHFIDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
||||| : ||||| : ||||| : ||||| : ||||| : |||||
a040      HGARHFIDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
70      80      90      100     110     120

130     140     150     160     170     180
m040.pep PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
||||| : ||||| : ||||| : ||||| : ||||| : |||||
a040      PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
130     140     150     160     170     180

190     200     210     220     230     240
m040.pep TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGT LAETLSAQEAQSLAEHAGQTRR
|| : ||||| : ||||| : ||||| : ||||| : |||||
a040      TFHLDMLQTAASVAVSLQAEKLVYLTLSDGISRPDGT LAVTLSAQEAQSLAEHAGGETRR

```

229

	190	200	210	220	230	240
			250	260	270	
m040.pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI					
a040	LISSAVAALIEGGVHRVQILNGAADGSLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI					
	250	260	270	280	290	300
	280	290	300	310	320	330
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFEADCGEIACLAVS					
a040	AALIRPLEEQGILLHRSREYLENHISEFSILEHDGNLYGCAALKTFEADCGEIACLAVS					
	310	320	330	340	350	360
	340	350	360	370	380	390
m040.pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY					
a040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY					
	370	380	390	400	410	420
	400	410				
m040.pep	RSNGRNSHILVRRLLHRX					
a040	RSNGRNSHILVRRLLHRX					
	430					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 040 shows 88.3% identity over a 436 aa overlap with a predicted ORF (ORF 040.ng) from *N. gonorrhoeae*:

m040/g040

m040.pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI	60
g040	MNAPDSFVAHFREAAPYIRQMRGTTLVAGIDGRLLLEGGTLNKLAADIGLLSQLGIRLVLI	60
m040.pep	HGARHFLLDRHAAQGRTPHYCRGLRVTDSETSLEQAQQFAGTVRSRFEAALCGSVSGFARA	120
g040	HGAYHFLLDRLAAQGRTPHYCRGLRVTDSETSLGQAQQFAGTVRSRFEAALCGSVSGFARA	120
m040.pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK	180
g040	PSVPLVSGNFLTARPMGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWMPPLGHSYSGK	180
m040.pep	TFYLDMLQTAASAAVSLQAELVYLTLSDGISRDPDGTLAETLSAQEAQSLAEHAGGQTRR	240
g040	TFNLDMVQAAASVAVSLQAELVYLTLSDGISRDPDGTLAETLSAQEAQSLAEHAASETRR	240
m040.pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI	276
g040	LISSAVAALIEGGVHRVQILNGAADGSLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI	300
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFEADCGEIACLAVS	336
g040	AALIRPLEEQGVLLHRSREYLENHISEFSILEHDGDLYGCAALKTFEADCGEIACLAVS	360
m040.pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY	396
g040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY	420
m040.pep	RSNGRNSHILVRRLLHRX	413
g040	RSNGRNPILVRRLLHRX	437

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 153>:

```
g041.seq
1   ATGAGTTCGC CCAAACACAT CGGCTTGCAG GCGGCGAGCA ACGGCGGCCT
51  GATTACCGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATC GGTGCGCTGG
101 TGTGCGAAGT ACCGCTGACC GATATGATCC GTTATCCGCT GCTGTCCGCC
151 GGTTCAGATT GGACGGACGA ATACGGCAAT CCGCAGAAAT ACGAAGCCTG
201 CAAACGCCGG CTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCACTC ATTAACACCA GCCTCAGCGA CGACCGCGTC
301 CATCCCGCCC ACGCGCTCAA ATTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCAACGCGA ATCCGCCGAC AAACCTGCCT GCGTCTTGCT GTTTTGTAAA
451 GAATTTTGG GATAA
```

This corresponds to the amino acid sequence <SEQ ID 154; ORF 041.ng>:

```
g041.pep
1   MSSPKHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQYEACKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQRESAD KLACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 155>:

```
m041.seq
1   ATCAGTTCGC CCGAACACAT CGGCTTGCAG GCGGCGAGCA ACGGCGGACT
51  GATTACTGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATC GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAGACT GGACAGACGA ATACGGCAAT CCGCAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCGGC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCAACGCGA ATCCGCCGAC GAACTCGCCT GCGTCTTGCT GTTTTGTAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 156; ORF 041>:

```
m041.pep
1   ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSQSW LYSPDGGGHT GNGTQRESAD ELACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 157>:

```
a041.seq
1   ATCAGTTCGC CCGAACACAT CGGCTTGCAG GCGGCGAGCA ACGGCGGACT
51  GATTACTGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATA GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAGACT GGACAGACGA ATACGGCAAT CCGCAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CGCAGCGCGA AGCCGCCGAC GAACTCGCCT GCGTCTTGCT GTTTTGTAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 158; ORF 041.a>:

```
a041.pep
1   ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQREAAD ELACVLLFLK
```

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
a041	ISSPEHIGLQGGSSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRVRHPAHALKFYAKLRETSQSW					
a041	PQKYEVCKRRLGELSPYHNLSGDIDYPPALITTSLSDDRVRHPAHALKFYAKLRETSQSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGGHTGNGTQRESADELACVLLFLKEFLGX					
a041	LYSPDGGGGHTGNGTQREADELACVLLFLKEFLGX					
	130	140	150			

m041/g041

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	:   :					
g041	MSSPKHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGELSPYHNLSDGIDYPPALITTSLSDDRVP	PAHALK	FYAKL	RETS	SAQSW	
		:				
g041	PQKYEACKRRLGELSPYHNLSDGIDYPPALITTSLSDDRVP	PAHALK	FYAKL	RETS	SPQSW	
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESAD	LACVLL	FLKEFLGX			
g041	LYSPDGGGHTGNGTQRESAD	LACVLL	FLKEFLGX			
	130	140	150			

g041-1.seq

1	ATGAAATCCT	ACCCCGACCC	CTACCGCCAT	TTTGA AAAC	TCGATTCGC
51	CGAAACGCAA	AACTTCGCTG	CTGAAGCAAT	TGCCGAAACG	CGCGCGCGTT
101	TTTTAAACAA	CGACAAGGCG	CGCGCACTTT	CAGACGGCAT	TTTGAATCAA
151	ATGCAGAGCA	CGCGGCAGAT	TCCGTTTGTG	CAGGAACACG	CGCGCGCGAT
201	GTACCATTTC	CATCAGAATG	CGGAATATCC	GAAGGCGGTG	TACCGCATGT
251	GTACGGCGCG	GACCTACCGT	TCCGGCTATC	CCGATGGAA	AATCCTGTGT
301	TCCGTGGCGG	ATTTCGATGA	TGTGCTCGGG	GACGATGTGT	ATTTGGGCGT
351	CGTGTGCGAC	TTGGTGGAGC	AGCCCAACCG	CGCGCTGCTG	ACTTTGAACA
401	AATCGGGCGG	CGATACGGCG	TATACGCTGG	AAGTGGATTT	GGAAGCAGGG
451	GAATTGGTAG	AGGGCGGTTT	TCACTTTCCG	GCAGGCAAAA	ACCATATGTTC
501	TGGCGCGCAT	GA AAACAGCT	TGTGGTTCGG	TCCGGCTTGG	GACGACACGC
551	AGTTGACCGA	ATCGGCTAT	CCGCGCGAAG	TGTGCTGTGT	GGAACGCGGG

```

601 AAGAGTTTCG AGGAAAGCCT GCCGGCGTAC CAAATCGATA AAGGCGCGAT
651 GATGGTAAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AGACGTATTT GCAGGTGTCG
751 TCCGAAGGCG GGGCGAAACC GTTGAACCTG CCTAATGATT GCGATGTGGT
801 CGGCTATCTG GCGGGACATC TTTTGCTGAC GCTGCGCAAG GACTGGCACC
851 GCGCGAACCA AAGCTATCCG AGTGGCGCGT TGGTGGCGGT GAAACTGAAT
901 CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCAAGCCTGC
1001 TGGAGAATGT ACAAGGCCGT CTGAAAGCGT GGCGGTTTGC CGACAGCAAA
1051 TGGCAGGAAG CCGAGTTGCC GCACCTGCCC TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGCGGCG ACGTGGTTTA TCTTGCCGCC AGCGATTTCa
1151 CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAAGTACC
1201 GTCATGCGCG TCCAGCCGCA GCAGTTTGTT TCAGACGGCA TCGAAGTGGC
1251 GCAGTTTGGG GCGGTGTCGT CCGACGGCGA ACGCATTCCT TATTTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACACGCCGA CCTTAGTCTA TGCTTACGGA
1351 GGTTCGGCA TTCTGAATT GCCGCATTAT CTGGGCAGCG TCGGCAAAATA
1401 TTGGCTGSAA GAGGGCAATG CCTTGTATT GGCAAAACATC CGCGGCGGCG
1451 GAGAATTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAC
1501 AAAAGCGTTG ATGATTTGTT GGCAGTCGTG CGTGATTGTG CCGAACGCGG
1551 CATGAGTTCG CCAAAACACA TCGGCTTGCA GGGCGGCAGC AACGCGCGGC
1601 TGATTACCGC CGCCGCCTTC GTGCGGAAC CGCAAAGCAT CGGTGCGCTG
1651 GTGTGCGAAG TACCGCTGAC CGATATGATC CGTTATCCGC TGCTGTCCGC
1701 CGGTTCAGT TGGACGGACG AATACGGCAA TCCGAGAAA TACGAAGCCT
1751 GCAAACGCCG GCTGGGCGAA TTGTGCGCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCACT CATTACCACC AGCCTCAGCG ACGACCGCGT
1851 CCATCCCGCC CACGCGCTCA AATTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CAAACTCGCC TCGGTGTGCG TGTTTTTGAA
2001 AGAATTTTGG GATAAA

```

This corresponds to the amino acid sequence <SEQ ID 160; ORF 041-1.ng>:

g041-1.pep

```

1 M KSYDPYR H FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILNQ
51 M QDTRQIPFC QEHRARMYHF HQNAEYPKGV YRMCTAATYR SGYPEWKILF
101 S VADFDLLG DDVYLGGVSH LVEQPNRALL TLNKSGGDTA YTLEVDLEAG
151 E LVEGGFHF P AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 K SFEEELPAY QIDKGAMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 S EGGAKPLNL PNDCDVVGYL AGHLLLTLRK DWHRANQSY S GALVAVKLN
301 R GELGAAQLL FAPDETQALE SVETTKRFV ASLLENVQGR LKAWRFADSK
351 W QEAELPHLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
401 V MRLQPQFV SDGIEVRQFW AVSSDGERIP YFHVGNKAP DTPTLVYAYG
451 F GFIPELPHY LGSVGVKYLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501 K SVDDLAVV RDLSEGMSS PKHIGLQGG NGGLITAAAF VREPQSIGAL
551 V CEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEACKRRLGE LSPYHNLSDG
601 I DYPPALITT SLSDRVHFA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 T QRESADKLA CVLLFLKEFL G*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 161>:

m041-1.seq

```

1 A TGAAATCCT ACCCGACCC CTACCGCCAT TTTGAAACC TCGATTCCGC
51 C GAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 T TTTAGAAAA CGACAAGGCG CGCGCGCTTT CAGACGGCAT TTTGGCGCAG
151 T TGCAGGACA CGCGGCAGAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
201 G TACCATTTC CATCAGGACG CGGAGTATCC GAAGGGCGTG TACCGCGGTG
251 G TACCGCGCG GACGTATCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
301 T CGGTGGCGG ATTTTCGACGA ATTGCTTGGC GACGATGTGT ATTTGGGCGG
351 C GTGTCGCAC TTGGTGGAA AGCCCAACCG CGCGTTGTTA AACTGAGCA
401 A ATTGGGCAG CGATACGGCG TACACGCTGG AAGTGGATTG GGAAGCAGGG
451 G AGTTGGTCG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 G TGGCGCGAT GAAAAACAGC GTGGGTGTG TCCGCTTGG AACGAACGCC
551 A GTTGACCCA ATCGGGCTAT CCGCGCGAAG TATGGCTGGT GGAACGCGGC
601 A AGAGTTTCG AGGAAAGCCT GCCTGTGTAT CAAATCGGCG AAGACGGCAT
651 G ATGGTGAAC GCGTGGCGTT ATCTCGATCC GCAGGGTTCG CCGATTGATT
701 T GATTGAAGC GTCGGACGGT TTTTACACCA AAACCTATTT GCGGCTCTCA

```



```

751 GCCGAAGGCG AGGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTGGT
801 CGGCTATCTG GCGGGGCATC TTTTGCTGAC GCTGCGCAAG GACTGGAACC
851 GCGCGAACCA AAGCTATCCG AGCGGCGCGC TGGTGGCGGT GAAGCTGAAT
901 CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAG CGACCAAGCG TTTTGTGGTG GCGAGCCTGT
1001 TGGAGAACGT ACAAGGCCGT CTGAAAGCAT GGCGGTTTGC CGACGGCAAA
1051 TGGCAGGAAG TCGAATTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCT TGGGGCGGCG ACGTGGTTTA CCTTGCCGCC AGCGATTTC A
1151 CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAAC TGAC
1201 GTCATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA
1251 GCAGTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCTT TATTTCACG
1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CCGGCTCTA TGCCTACGGC
1351 GGTTCGGCA TTCCGAAATT GCCGATTAT CTGGCAGCA TTGGCAAATA
1401 TTGGCTGGAA GAGGGCAATG CCTTGTATT GGCGAACATC CGCGCGCGG
1451 GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTATT GGCAGTCGTG CGCGATTTGT CCGAACGCGG
1551 TATCAGTTTC CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601 TGATTACTGC CGCCGCCCTT GTGCGCGAAC CGCAAAGCAT CGCGCGCGTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAACGCGG GTTGGGCGAA TTGTGCGCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCCG
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CGAACTCGCC TCGCTCTTGC TGTTTTGA
2001 AGAGTTTTTG GGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 162; ORF 041-1>:

**m041-1.pep**

```

1 MKSYDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQ
51 LQDTRQIPFCQEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILF
101 SVADFDELLGDDVYLGGVSHLVEQPNRALLTL SKLGSDTA YTLEVDLEAG
151 ELVEGGFHFAPAGKNHVSWRDENSVWVCPAWNERQLTQSGYPREVWLVERG
201 KSFEESLPVYQIGEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLRVS
251 AEGEAKPLNL PND CDVVGYL AGHLLTLRK DWN RANQSYPSGALVAVKLN
301 RGELGAAQLL FAPDETQALE SVETTKRFVV ASLENVQGR LKAWRFADGK
351 WQEVLPRLP SGALEMTDQPWGGDVVYLAA SFTTPTLTLF ALDLNVMELT
401 VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGNKNAAP DMPTLVYAYG
451 GFGIPELPHY LGSIGKYWLE EGNFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLAVV RDLSEGISS PEHIGLQGG NGGLITAAAF VREPOSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDRRVHPA HALKFYAKLR ETSAQSWLYS PDGGGHTGNG
651 TQRESADELA CVLLFLKEFL G*

```

**m041-1/g041-1** 94.6% identity in 671 aa overlap

	10	20	30	40	50	60
m041-1.pep	MKSYDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQ LQDTRQIPFC					
g041-1	MKSYDPYRHFENLDSAETQNFAAEANAETRARFLNNDKARALSDGILNQMQDTRQIPFC					
	10	20	30	40	50	60
m041-1.pep	QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH					
g041-1	QEHRARMYHFHQDAEYPKGVYRMCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH					
	70	80	90	100	110	120
m041-1.pep	LVEQPNRALLTL SKLGSDTAYTLEVDLEAGELVEGGFHFAPAGKNHVSWRDENSVWVCPAW					
g041-1	LVEQPNRALLTLNKS GGDYATY TLEVDLEAGELVEGGFHFAPAGKNHVSWRDENSVWVCPAW					
	130	140	150	160	170	180
m041-1.pep	NERQLTQSGYPREVWLVERGKSFEESLPVYQIGEDGMMVN AWRYLDPQGS PIDLIEASDG					
g041-1	NERQLTQSGYPREVWLVERGKSFEESLPVYQIGEDGMMVN AWRYLDPQGS PIDLIEASDG					

sp|P55577|Y4NA\_RHISN PROBABLE PEPTIDASE Y4NA >gi|2182536 (AE000086) Y4nA [Rhizobium sp.  
NGR234] Length = 726

Identities = 217/682 (31%), Positives = 331/682 (47%), Gaps = 22/682 (3%)

Sbjct: 102 FARDGMIDNFWQDGHVQGLWRRRTTWESYRSGNPQWRTILDVDALSKEGKTVWFEGGDC 161

Query: 121 LVEQPNRALLTLKLGSDTAYTLEVDLEAGELVEGGFHFPAKGNNHWSKRDENS VWVCPAW 180  
L N L+ LS G D E D+ GE V+ GF P GK V+W DEN+++V W

Sbjct: 162 LPPTSNLCLIRLSGGKADADVREFDIAKGEFVKEGFVLPEGKQSVTWVDENTIIYVTREW 221

Query: 181 NERQLTQSGYPREVWLVERGKSFEESLPVYQ-----IGEDGMM--VNAWRYLDPQGSPI 232  
 ++T SGY +V+RG+S ++++ +++ E G++ ++ -D +

Sbjct: 222 TPGEVTSYGAYVTKVVRGQSLDQAVEIFRGQKKDVS AERGVLRDIDGKYVMDTSYRGL 281

Query: 233 DLIEASDGFYTKTYLRVSAEGEAKPLNLPNDCDVVGYLAGHLLTLRKDNRRANQS-YPS 291  
 D FY + + L LP GY G + L+ DW A + + +

Sbjct: 282 DFFNTELAIFYPNHG----PDTRKVVLPPTTAVFSGYYKGQAIYWLKSDWTS AKGTVFHN 337

Query: 292 GALVAVKLNREGELGAAQL----LFAPDETOALESVETTKRFVVASLLENVQGRKAWRFA 347  
 GA++A L A++ LF P+E Q++ TK +V S+L NV ++++ F

Sbjct: 338 GAIIFDLKAAALADPARVEPLVLFMPNEHQSVAGTTQTKNPLVLSEILNVTSEVRSFDFG 397

Query: 348 DGKWQEVLPPLPSGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMRRQPQ 407  
 G W +L + L +T D +++ + F P TLF D ++ + P

Sbjct: 398 KGGWSSFKLALPENSTLSLTSSDDESDQLFVFESEGFLEPSTLFCADAATGQVEKITSTPA 457

Query: 408 QFDSGGINVQQFWTTSADGERIPYFHUGANAAP---DMPTLVYAYGGFGIPELPHYLGS 464  
 +FD+ G+ QQFW TS DG ++PYF V + PT++YAYGGF IP P Y +

Sbjct: 458 RFDAGGLQAQQFWATSKDGTKVYPFLVARKDKVLDGNTPTILYAYGGFQIPMOPSYSAVL 517

Query: 465 GKYWLEEGNAFVLNIRGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSEGGISSPEHI 524  
 GK WLE+G A+ LANIRGGGEFGP+WH A ++ + DD AV +DL + ++S H+

Sbjct: 518 GKLWLEKGGAYALANIRGGGEFGPKWHDAGLKTNRQRYDDFQAVAQDLIAKKVTSTPHL 577

Query: 525 GLQGGSSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEV 584  
 G+ GGSNGGL+ ++ P A+V +VPL DM+ + +SAG+SW EYG+P V

Sbjct: 578 GIMGGSNGGLMGVQMTQRPDLWNAVVIQVPLDMVNFTRMSAGASWQAEYGSPPD-PVE 636

Query: 585 KRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAAHALKFYAKIRETSAQSWLYSPDGG 644  
 L +SPYHN+ G+ YP TS DDRV P HA K A + + Y G

Sbjct: 637 GAFLRISIPYHNKAGVAYPEPFETSTKDDRVGPVHARKMAALFEDMGLPFYYENIEG 696

Query: 645 GHTGNGTQRESADELACVLLFL 666  
 GH +E A A +++

Sbjct: 697 TAAAAANLQEHARLALLEYIYM 718

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 163>:

a041-1.seq

```

1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAAACAA CGACAAGGCA CGCGCATTTG CTGACGGCAT TTTGGCGCAG
151 TTGCAGGACA CGCGGCAAAT TCCGTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAAGATG CGGAATATCC GAAAGGCGTG TACCGCGTGT
251 GTACCGCGGC GACTTACCGT TCGGGCTATC CTGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTGACGCA ATTGCTCGGT GACGATGTAT ATCTAGGCGG
351 CGTGTCGCAC CTGGTGGAAAC AGCCCAACCG CGCGTTGTTA AACTGAGCA
401 AATCGGGCGG CGATACCGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAGTTGGTAG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
551 AGTTGACCGA ATCGGGCTAT CCGCGCGAGG TGTGGCTGGT GGAACGCGGC
601 AAGAGTTTCG AGGAAAGCCT GCCGGTGTAC CAAATTGCTG AAGACGGCAT
651 GATGGTGAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTTC CCGATTGATT
701 TGATTGAACG GTCTGACGGT TTTTACACCA AAACCTATTT GCAGGTCTCA
751 GCCGAAGGCG AAGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTAGT
801 CGGCTATCTG GCCGGACATC TTTTGTGAC CTTGCGTAAA GACTGGCACC
851 GCGCGAACCA AAGCTATCCG AGTGGCGCAT TGGTAGCAGT AAAATTAAC
901 CGCGGCGAAT TGGGCGCGGC GCAGCTTTG TTTGCGCCCA ATGAACGCA
951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTCTGT GCGAGCCTGC
1001 TGGAAAACGT ACAGGGTCGT CTGAAAGCGT GCGCTTTTAC TGATGGCAAA
1051 TGGCAGGAAA CCGAGTTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGGGGCG ACGTAGTTTA CCTTGCCGCC AGCGATTCTA
1151 CCACGCCGCT GACGCTGTTT GCATTGGATT TGAACGTGAT GGAACGTACC
1201 GTCATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA
1251 GCAGTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCACG
1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGTGCTCTA TGCCTACGGC
1351 GGTTCGGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAAATA
1401 TTGCTGGAA GAGGGCAATG CCTTTGTATT GCGAACATC CGCGCGCGCG

```

```

1451 GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTTATT GGCAGTCGTG AGCGATTTGT CCGAACGCGG
1551 TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601 TGATTACTGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT AGGCGCGCTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAACGCGG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCGCGC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACGCAGCGCG AAGCCGCCGA CGAACTCGCC TCGGTGTTGC TGT'TTTTGA
2001 AGAGTTTTTG GGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 164; ORF 041-1.a>:

**a041-1.pep**

```

1  MKSYDPYRHFENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILAQ
51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKSGGDTA YTLEVDLEAG
151 ELVEGGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPVY QIAEDGMMVN AWRYLDPQGS FIDLIEASDG FYTKTYLQVS
251 AEGEAKPLNL PNDCDVVGYL AGHLLTLRK DWHRANQSY SGALVAVKLN
301 RGELGAAQLL FAPNETQALE SVETTKRFV ASLLENVQGR LKAWRFTDGK
351 WQETELPRLP SGALMTDQP WGGDVVYLAA SDFTTPTLTF ALDLNVMELT
401 VMRRPQQQFD SDGINVQQFW TTSADGERIP YFHVGNKNAAP DMPTLVYAYG
451 GFGIPELPHY LGSIGKYWLE EGNFVLANI RGGGEFGPRW HQAAOGISKH
501 KSVDDLLAVV SDLSEGISS PEHIGLQGGG NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDRVHFA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQREAADELA CVLLFLKEFL G*

```

**a041-1/m041-1** 97.9% identity in 671 aa overlap

	10	20	30	40	50	60
a041-1.pep	MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLNNDKARALSDGILAQ LQDTRQIPFC					
m041-1	MKSYPDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQ LQDTRQIPFC					
	10	20	30	40	50	60
	70	80	90	100	110	120
a041-1.pep	QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVVYLGGVSH					
m041-1	QEHRARMYHFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVVYLGGVSH					
	70	80	90	100	110	120
	130	140	150	160	170	180
a041-1.pep	LVEQPNRALLTLSKSGGDTAYTLEVDLEAGELVEGGGFHFPAGKNHVSWRDENSVWVCPAW					
m041-1	LVEQPNRALLTLSKLGS DTA YTLEVDLEAGELVEGGGFHFPAGKNHVSWRDENSVWVCPAW					
	130	140	150	160	170	180
	190	200	210	220	230	240
a041-1.pep	DERQLTESGYPREVWLVERGKSFEESLPVYQIAEDGMMVN AWRYLDPQGS FIDLIEASDG					
m041-1	NERQLTQSGYPREVWLVERGKSFEESLPVYQIGEDGMMVN AWRYLDPQGS FIDLIEASDG					
	190	200	210	220	230	240
	250	260	270	280	290	300
a041-1.pep	FYTKTYLQVSAEAEAKPLNPNDCDVVGYLAGHLLTLRKDWHRANQSYPSGALVAVKLN					
m041-1	FYTKTYLRVSAEAEAKPLNPNDCDVVGYLAGHLLTLRKDWNRANQSYPSGALVAVKLN					
	250	260	270	280	290	300
	310	320	330	340	350	360
a041-1.pep	RGELGAAQLLFAPNETQALESVETTKRFVVASLLENVQGR LKAWRFTDGK WQETELPRLP					
m041-1	RGELGAAQLLFAPDETQALESVETTKRFVVASLLENVQGR LKAWRFADGK WQEVLPRLP					
	310	320	330	340	350	360

a041-1.pep	370	380	390	400	410	420
	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSGGINVQQFW					
m041-1	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSGGINVQQFW					
	370	380	390	400	410	420
a041-1.pep	430	440	450	460	470	480
	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIKGYWLEEGNAFVLANI					
m041-1	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSIKGYWLEEGNAFVLANI					
	430	440	450	460	470	480
a041-1.pep	490	500	510	520	530	540
	RGGGEFGPRWHQAAQGISKHKSVDDLAVVSDLSERGISSPEHIGLQGGSNGLITAAAF					
m041-1	RGGGEFGPRWHQAAQGISKHKSVDDLAVVSDLSERGISSPEHIGLQGGSNGLITAAAF					
	490	500	510	520	530	540
a041-1.pep	550	560	570	580	590	600
	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
m041-1	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
	550	560	570	580	590	600
a041-1.pep	610	620	630	640	650	660
	IDYPPALITTSLSDDRVHPAHALKFYAKLRETSQSWLYSPDGGGHTGNGTQREAADELA					
m041-1	IDYPPALITTSLSDDRVHPAHALKFYAKLRETSQSWLYSPDGGGHTGNGTQRESADELA					
	610	620	630	640	650	660
a041-1.pep	670					
	CVLLFLKEFLGX					
m041-1	CVLLFLKEFLGX					
	670					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 165>:

**g042.seq**

```

1  ATGACGATGA TTTGCTTGCG CTTCCAagcG TTCGTGCCGC ATACCAGCGC
51  GTTATCCAAC ACTTCCACGG CAGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TGCGGTTCGAT GATGAAAATC CAGCCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACGG GCTGCCCGTG CCCTTCGTTG CGTAAAGATT CGTCCACGGG
201 CGGCAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GATTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCTGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG cggCTTCGCG CTTTGGGCG AACAGCGCGT CAATCTGCGC
351 ATTCAATTCC GCCACGCGCG CTCCTTACC GAAAATCCGC GACAGGGTCT
401 CCATCTGCTT CTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAAA
451 TCTATGgtgG TCGCGTTTTT CGCCAACGTG TCATACGCTT CCGCACCCGG
501 CCGCCCGGTA ATGACAAACT GCGGATTGTG GCGGTGCAGG GATTCGCAAT
551 CGGGCTCAAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAAA TGCAGGCTGC
601 AAATAG
```

This corresponds to the amino acid sequence <SEQ ID 166; ORF 042.ng>:

**g042.pep**

```

1  MTMICLRFOA FVPHTSALSN TSTAAGPSCP MAAVRSMMKI QPGFFSLMYS
51  KETGCPGPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSSTPRP
101 LPLAASRFAW NSASICAFNS ATRASLPKIR DRVSICFSPV VRLLPLSTVK
151 SMVVAFFANC SYASAPGPPV MTNCGLWRCR DSQSGSNSVP TVAALSNAAG
201 K*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 167>:

## m042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAmT ACTTCGACAG CCGcCGGCCy TTCyTGCCCG ATGGCGGCCG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTTC
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAAcCG ACAGCACCGC CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAAATCCG GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTAAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 168; ORF 042>:

## m042.pep

```

1 MTMICLRFQA FVPRTSALSX TSTAAGXSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPGPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSSTPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTSXGLXRCR ASXSGSNSVP TVAALSNAGC
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 169>:

## a042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCCG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTTC
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAAACC ACAGCACCGC CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAAATCCG GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCCAACGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 170; ORF 042.a>:

## a042.pep

```

1 MTMICLRFQA FVPRTSALSX TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPGPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSSTPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*GL*RCR AS*SGSNSVP TVAALSNAGC
201 K*

```

m042/a042 99.0% identity over a 201 aa overlap

	10	20	30	40	50	60
m042.pep	MTMICLRFQAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPGPSL					
a042	MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPGPSL					
	10	20	30	40	50	60
m042.pep	70	80	90	100	110	120
	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSSTPRPLPLAASRVWANSASICAFNS					
a042	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSSTPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
m042.pep	130	140	150	160	170	180
	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					

Homology with a predicted ORF from *N. gonorrhoeae*

m042/q042

	10	20	30	40	50	60
m042.pep	MTMICLRFOAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
	:					
g042	MTMICLRFOAFVPHTSALSNTSTAAGSPCPMAAVRSMMKIQPGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m042.pep	RKDSSTGGRMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS					
g042	RKDSSTGGRMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRFWANSASICAFNS					
	70	80	90	100	110	120
	130	140	150	160	170	180
m042.pep	AARASLSPKIRAKVSI CF SPLVRILPLSTVRS MVVAFFANC SYASAPGPVMT SXGLXR CR					
	:					
g042	ATRASLSPKIRDRVSI CF SPLVRILPLSTV KSMVVAFFANC SYASAPGPVMT NCGLWR CR					
	130	140	150	160	170	180
	190	200				
m042.pep	ASXSGSNSVPTVAALSNAGCKX					
g042	DSQSGSNSVPTVAALSNAGCKX					
	190	200				

m042-1.seq

1	ATGACGATGA	TTTGCTTGCG	CTTCCAAGCG	TTCGTGCCGC	GTACCAGCGC
51	GTTATCCAAT	ACTTCGACAG	CCGCCGGCCC	TTCCTGCCCG	ATGGCGGCGG
101	TACGCTCGAT	GATGAAAATC	CAATCGGGGT	TTTTCTCTTT	GATGTATTTC
151	AAGGAAACAG	GCTGCCCGTG	CCCCTCGTTT	CGTAAAGATT	CGTCTACAGG
201	CGGTAGGCCG	ATGTGCGCGT	CGATCCAACT	TGCCAACGCG	GACTCGCTGC
251	CGAAGCGCGA	CACCTTGTTG	CCCGTAACCG	ACAGCACCAG	CCCGCGTCTC
301	TTGCCTTTGG	CGGCTTCGCG	CGTTTGGGCG	AACAGCGCGT	CAATCTGCGC
351	CTTCAATTCC	GCCGCGCGCG	CTTCCTTGCC	GAAAATCCGC	GCCAAGGTCT
401	CCATCTGCTT	TTCGCCGCTG	GTGCGGATAT	TGCCGTTGTC	CACCGTCAGA
451	TCTATGGTTG	TCGCGTTTTT	CGTAACTGT	TCATACGCTT	CCGCGCCCGG
501	CCCGCCGGTA	A			

m042-1.ppt

1 MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS  
51 KETGCGCPSL RKDSSTGGRP MSPCIQLANR DCPVKADTLL PVTDTSTSRP  
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR  
151 SMVVAFFANC SYASAPGPPV MTS\*

m042-1/g042 95.4% identity in 173 aa overlap

240

	10	20	30	40	50	60
m042-1.pep	MTMICLRFAQFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
g042	MTMICLRFAQFVPHTSALSNTSTAAGPSCPMAAVRSMMKIQPGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m042-1.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
g042	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
	130	140	150	160	170	
m042-1.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSX					
g042	ATRASLPKIRDRVSICFSPLVRILPLSTVKSMVVAFFANC SYASAPGPPVMTNCGLWRCR					
	130	140	150	160	170	180
g042	DSQSGSNSVPTVAALS NAGCKX					
	190	200				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 173>:

```

a042-1.seq
1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51  GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTTC
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAT TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTCCTTGCC GAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCGGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCCAATGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA A

```

This corresponds to the amino acid sequence <SEQ ID 174; ORF 042-1.a>:

```

a042-1.pep
1  MTMICLRFAQ FVPRTSALSNTSTAAGPSCPMAAVRSMMKI QSGFFSLMYS
51  KETGCPCPSL RKDSSTGGRP MSPCIQLANRDCVPKADTLL PVT DSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*

```

m042-1/a042-1 100.0% identity in 173 aa overlap

	10	20	30	40	50	60
m042-1.pep	MTMICLRFAQFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
a042-1	MTMICLRFAQFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m042-1.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
a042-1	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
	130	140	150	160	170	
m042-1.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSX					
a042-1	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSX					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 175>:



```

g043.seq
1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GTCGGCCCAT CAGCACTTTT
51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GCCTGAACGC TTCgttgaAC
101 CGTCCCGCGT ggcggtagcc gcAAAAGTGC ATcGCGGCTT GGATGGTGCT
151 GCCCGATTTCG ATGAGGGcga gcGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 GTCCGGCGAC GGTTCGCGCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTTCGTTC GCGCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGGCG
301 GGCGAATTCG CTGTTCAAAA TATCGGCGGC TTCGTCTATG CGCCGGCGGC
351 GGTAGCCGTT GTCGTGGCGG CGGAAGGTGA AGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 176; ORF 043.ng>:

```

g043.pep
1  MVVSNQNIYA VGPSALFHIR RQKSVMPPER FVEPSRVAVA AKVHRGLDGA
51  ARFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQFDA AGDFGDGQRA
101 GEFVQNIGG FVYAPAAVAV VVAEAGEA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 177>:

```

m043.seq
1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCTT CAGCACTTCT
51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101 CGTCCCGCGT GGCGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT
151 GCCGATTTCG ATGAGGGCGA GCGCGTGTTT CAGCCGCAGG CGGCGCAqGC
201 ATCCGGCGAC GGTTCGCGCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTTCGTTC GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301 GGCGAATTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351 GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 178; ORF 043>:

```

m043.pep
1  MVVSNQNIYA AGPSALLHIR RQKSVMPSEF FVEPSRVAVA AKVHGGLDGA
51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAEAGEAQ*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* *menA* with *menB*

ORF 043 shows 89.8% identity over a 128 aa overlap with a predicted ORF (ORF043.a) from *N. gonorrhoeae*:

**m043/g043**

	10	20	30	40	50	60
m043.pep	MVVSNQNIYAAGPSALLHIRRQKSVMPSEFVEPSRVAVAAKVHGGLDGAAGFDEGERVF					
	:    :     :     :     :     :     :     :     :     :					
g043	MVVSNQNIYAVGPSALFHIRRQKSVMPPEFVEPSRVAVAAKVHGLDGAARFDEGERVF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m043.pep	QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV					
	:     :     :     :     :     :     :     :     :					
g043	QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVQVAFVQSDAAGDFGDGQRT					
	70	80	90	100	110	120
	130					
m043.pep	VVAEAGEAQX					
g043	VVAEAGEAAX					
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 179>:

```

a043.seq
1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCTT CAGCACTTCT
51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101 CGTCCCGCGT GGCGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT

```

242

```

151 GCCGGATTTC ATGAGGGGGA GCGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 ATCCGGCGAC GGTTTCGCCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTCGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301 GCGGAATTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351 GGTAAACGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 180; ORF 043.a>:

```

a043.pep
1  MVVSNQNIYA AGPSALLHIR RQKSVMPSER FVEPSRVAVA AKVHGGLDGA
51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

```

m043/a043 100.0% identity in 129 aa overlap

```

              10      20      30      40      50      60
m043.pep      MVVSNQNIYAAGPSALLHIRRQKSVMPSERFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
              |||||
a043           MVVSNQNIYAAGPSALLHIRRQKSVMPSERFVEPSRVAVAAKVHGGLDGAAGFDEGERVF
              10      20      30      40      50      60

              70      80      90      100     110     120
m043.pep      QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
              |||||
a043           QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFGDGQRTGEFVLQDVGGFVYAPTAVTV
              70      80      90      100     110     120

              130
m043.pep      VVAAEGEAQX
              |||||
a043           VVAAEGEAQX
              130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 181>:

```

g044.seq
1  ATGCTGCCCG ACCAGAGCGT CGAGTTCTTG CCACAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GGCGGCGGTT TTCCAGCCGT TCGGCTTCCA ACCGTGTATC
101 CAGTTTTCCTA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTAGCCCC TTCGATAACG GCGGTCAGCT
201 CCATGCGGTT GTTGGTGTTT TCGGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 CGGCTGCCGT AGCGCATTA

```

This corresponds to the amino acid sequence <SEQ ID 182; ORF 044.ng>:

```

g044.pep
1  MLPDQSVEFL PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
51  GAAAFERFQP FDNQGQLHAV VGGLRFAAEK FFFAAVAH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 183>:

```

m044.seq
1  ATGCCGTCCG ACTAGAGCGT CGAGTTCTTT CCAGAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GGAGGCGGTT TTCCAGCCGT TCGGCTTCCA ACCGTGTATC
101 CAGTTTTCCTA TGCCATTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTAGCTCC TTCGATGACG GCAGTCAGTT
201 CCATGCGGTT GTTGGTGTTT TCGGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 TGGCTACCGT AGCGCAyTAa

```

This corresponds to the amino acid sequence <SEQ ID 184; ORF 044>:

```

m044.pep
1  MPSDXSVEFF PEVVVFDGLF GGGFPAVALP TVYPVFHAI F DVLRVGADDD
51  GAAAFERFQS FDDGSQFHAV VGGLRFAAEK FFFVATVAH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 185>:

```
a044.seq
1  GTGCCGTCCG ACCAGCGCGT CGAGTTCTTT CCACAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GCGGCGCGTT TTCCAGCCGT TCGCTTCCA ACCGTGTATC
101 CAGTTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTCAGTCC TTCGATGACG GCGGTCAGTT
201 CCATACGGTT GTTGGTGGTT TCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 TGGCTGCCGT AGCGCATTA
```

This corresponds to the amino acid sequence <SEQ ID 186; ORF 044.a>:

```
a044.pep
1  VPSDQRVEFF PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
51  GAAAFERFQS FDDGGQFHTV VGGLRFAAEK FFFVAVAH*
```

m044/a044 91.0% identity over a 89 aa overlap

```

      10      20      30      40      50      60
m044.pep  MPSDXSVEFFPEVVVFDGLFGGGFPVALPTVYPVFHAI FDLRVGADDDGAAAFERFQS
          :||| :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
a044      VPSDQRVEFFPQVVVFDGLFGGGFPVALPTVYPVFHAVFDVLRVGADDDGAAAFERFQS
          10      20      30      40      50      60

      70      80      90
m044.pep  FDDGSQFHAVVGGLRFAAEKFFVATVAHX
          |||:|:|:||||:||||:||||:||||:||||
a044      FDDGGQFHTVVGGLRFAAEKFFVAAVAHX
          70      80      90
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 044 shows 86.5% identity over a 89 aa overlap with a predicted ORF (ORF 044.ng) from *N. gonorrhoeae*:

m044/g044

```

      10      20      30      40      50      60
m044.pep  MPSDXSVEFFPEVVVFDGLFGGGFPVALPTVYPVFHAI FDLRVGADDDGAAAFERFQS
          | | |||:|:||||:||||:||||:||||:||||:||||:||||:||||:||||
g044      MLPDQSVEFLPQVVVFDGLFGGGFPVALPTVYPVFHAVFDVLRVGADDDGAAAFERFQP
          10      20      30      40      50      60

      70      80      90
m044.pep  FDDGSQFHAVVGGLRFAAEKFFVATVAHX
          |:|:|:||||:||||:||||:||||:||||
g044      FDNGGQLHAVVGGLRFAAEKFFFAA VAHX
          70      80      90
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 187>:

```
g046.seq
1  ATGTCGGCAA TGCTGCGTCC GACAAGCAGC CCGCCGCGcc gCGCCTGTAT
51  GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
151 CTGATGGTTT CGGTATAGCC gaATATGGAA AGGCTGCCGt TTTcGTTGTT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TtcgctGGAA CGGACGCGCG
251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGagaGCGCG AGcagcaagt cggcatcttC
351 CgcgccggcG Cgttataatg tgAAGGGGGA TGCgccgttg ccgaAAACGG
401 TTTGAcatc gaggcggctg CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAAcgg TTACGTCCTT GTTGGTGATG GCGGCAAGGT TTTGCGCGAC
```

244

501 GGTAGAACCT ACCTGCCCGT TGCCTAAAAT GAGGATTTTC ACGGTATGGG  
551 TCGCCGGGTG A

This corresponds to the amino acid sequence <SEQ ID 188; ORF 046.ng>:

g046.pep  
1 MSAMLRPTSS PRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG  
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT  
101 MLVSSLRESA SSKSASSAPA RYNVKGDAPL PKTVWTSRRL PVSCNAFSSM  
151 SITVTSLLVM AARFCATVEP TCPLPKMRIF TVWVAG\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 189>:

m046.seq  
1 ATGTCGGCAA TGCTGCGTCC GACAAGCAst CCGC.r.sGC gCGcCTGTAT  
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCCG  
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCCGGG  
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTGCTTGT  
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG  
251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG  
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC  
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TCGCCCGTTG CCGAAAACGG  
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG  
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC  
501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG  
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 190; ORF 046>:

m046.pep  
1 MSAMLRPTSX PXXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG  
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT  
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM  
151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 191>:

a046.seq  
1 ATGTCGGCAA TGCTGCGTCC GACAAGCAGT CCGCCGCGCC GCGCCTGTAT  
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCCG  
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCCGGG  
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTGCTTGT  
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG  
251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG  
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC  
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TCGCCCGTTG CCGAAAACGG  
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG  
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC  
501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG  
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 192; ORF 046.a>:

a046.pep  
1 MSAMLRPTSS PRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG  
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT  
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM  
151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE\*

m046/a046 98.4% identity over a 186 aa overlap

	10	20	30	40	50	60
m046.pep	MSAMLRPTSXPXXRACMMTIRTRSSAKRKT CNAPGQSIRP ASCSVTSCSG LMVSVMPNME					
a046	MSAMLRPTSSPPRRACMMTIRTRSSAKRKT CNAPGQSIRP ASCSVTSCSG LMVSVMPNME					

245

	10	20	30	40	50	60
	70	80	90	100	110	120
m046.pep	RLPFSLFSSSLGLRYSRYSLETRAMRPGMLNRSAA TLQATMLVSSSLRESASSKSASSAPA					
a046	RLPFSLFSSSLGLRYSRYSLETRAMRPGMLNRSAA TLQATMLVSSSLRESASSKSASSAPA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m046.pep	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
a046	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
	130	140	150	160	170	180
m046.pep	TVWVAEX					
a046	TVWVAEX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 046 shows 97.3% identity over a 185 aa overlap with a predicted ORF (ORF 046.ng) from *N. gonorrhoeae*:

m046/g046

	10	20	30	40	50	60
m046.pep	MSAMLRPTSPXPRACMMTIRTRSSAKRKTCNAPGQSIRPASC SVTSCSGLMVSVMPNME					
g046	MSAMLRPTSSPPRRACMMTIRTRSSAKRKTCNAPGQSIRPASC SVTSCSGLMVSVMPNME					
	10	20	30	40	50	60
	70	80	90	100	110	120
m046.pep	RLPFSLFSSSLGLRYSRYSLETRAMRPGMLNRSAA TLQATMLVSSSLRESASSKSASSAPA					
g046	RLPFSLFSSSLGLRYSRYSLETRAMRPGMLNRSAA TLQATMLVSSSLRESASSKSASSAPA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m046.pep	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
g046	RYNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLVMAARFCATVEPTCPLPKMRIF					
	130	140	150	160	170	180
m046.pep	TVWVAEX					
g046	TVWVAGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 193>:

g047.seq

1	ATGGTCATCA	TACAGGCGcg	gcGCGGCGGG	CTGCTTGTCG	GACGCAGCAT
51	TGCCGACATC	GCCCAAGATT	TGCCCGACGG	GGCCGACTGC	CAAATCTGCG
101	CCGTTTACCG	CAACAACCGC	CTCATC	TCC	CCGCGCCGCA
151	ATCGAAGGCG	ACGAAATCCT	GTTTGCCGCC	GCCGCCGAAA	ACATCGGGGC
201	GGTCATACcc	gaATTGCGCC	CCAAAGAAAC	CAGCACC CGC	CGCATCATGA
251	TTGCCGGCGG	CGGCAACATc	tgetACCGCC	TCGCCAAGCA	GCTCGAACAC

246

```

301 GCATAcaacG TCAAAATCAT CGAATGCCGG CCGCGCcgTg ccgaATGGAT
351 AGCCGAAAAAC ctcgAcaaCA CCCTCGTCCT GCAAGGTTCTG Gcaaccgacg
401 aAaccctgct cgAcaacgaa tacatcgacg aaatcgaCGT ATTCTGCGCC
451 CTGACCAACG ACGACGAAAG CAACATTatg tCCGCCCTTT TGGCGAAAAA
501 CCTcgggcgCG AAGCgcgtca tcggCATCGT CAACCGCTCA AGCTACGTCTG
551 ATTTGCTCGA AGGCAACAAA ATCGACATCG TCGTCTCCCC CCACCTCATC
601 ACCATCGGCT CGATACTCGC CCACATCCG GCGGCGGACA TCGTTGCCGT
651 CCACCCATC CCGCGCGGCA CCGCGGAAGC CATCGAAGTC GTCGCGCACG
701 GCGACAAAAA AACTTCCGCC ATCATCGGCA GCGCATCAG CCGCATCAAA
751 TGGCCCGAAG GCTGCCACAT TGCCGCCGTC GTCCGCGCCG GAACCGGCGA
801 AACCATTATG GGACACCATA CCGAAACCGT CATCCAAGAC GGTGACCACA
851 TCATCTTTT CGTCTCGCGC CCGCGCATCC TGAACGAACT GGAGAAACTC
901 ATCCAAGTCA AAATGGGCTT TTTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 194; ORF 047.ng>:

**g047.pep**

```

1  MVIIQARRGG LLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGNI CYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCF
151 LTNDDESIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEEIEV VAHGDKKTS A IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 195>:

**m047.seq**

```

1  ATGGTCATCA TACAGgCGcG C..syGCGGA sTGCTTGTCG GACGCAGCAT
51  TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCCTCATC
151 ATCGAAGGCG ACGAAATCCT ATTTGCCGCC GCCGCCGAAA ACATCGGCGC
201 GGTCATACCC GAATTGCGCC CCAAAGAAAC CCAAAGAAAC CAGCcCmgmm
251 GcATCATGAT TkCCGGCGCG GGCAACATCG GCTACCGTCT CGCCAAGCAG
301 CTCGAACACG CATAACAAGT yAAAATCATC GAATGCCGGC CCGCGCGTGC
351 CGAATGGATA GCCGAAAACC TCGACAACAC CCTCGTCyTG CAAGGTTTCG
401 CAACCGACGA AACCTTGCTC GACAACGAAT ACATCGACGA AATCGACGTA
451 TTCTGCGCCC TGACCAACGA CGACGAAAGC AACATTATGT CCGCCCTTTT
501 GGCGAaAAAC CTCGGCGCGA AGCGCGTCAT CGGCATCGTC AACCGTCAA
551 GCTACGTCGA TTTGCTCGAA GGCAACAAAA TCGACATCGT CGTCTCCCCC
601 CACCTCATCA CCATCGGCTC GATACTCGCC CACATCCGGC GCGGCGACAT
651 CGTTGCCGTC CACCCCATCC GGCGCGGCAC GGCGGAAGCC ATCGAAGTCG
701 TCGCACACGG CGACAAAAAA ACTTCCGCCA TCATCGGCAG GCGCATCAGC
751 GGCATCAAAT GGCCCGAAGG CTGCCACATT GCCGCCGTCG TCCGCGCCGG
801 AACCGGCGAA ACCATTATGG GACACCATAC CGAAACCGTC ATCCAAGACG
851 GCGACCACAT CATCTTTTTC GTCTCGCGCC GGCGCATCCT GAACGAACTG
901 GAAAAACTCA TCCAGGTCAA AATGGGCTTT TTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 196; ORF 047>:

**m047.pep**

```

1  MVIIQARXXG XLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETQRN QPXXIMIXGG GNIGYRLAKQ
101 LEHAYNVKII ECRPRRAEWI AENLDNTLV LQGSATDETL DNEYIDEIDV
151 FCALTNDDES NIMSALLAKN LGAKRVIGIV NRSSYVDLLE GNKIDIVVSP
201 HLITIGSILA HIRRGDIVAV HPIRRGTAE A IEVVAHGDKK TSAIIIGRRIS
251 GIKWPEGCHI AAVVRAGTGE TIMGHHTETV IQDGDHIIFF VSRRIILNEL
301 EKLIQVKMGF FG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 197>:

**a047.seq**

```

1  ATGGTCATCA TACAGGCGCG GCGCGGCGGA CTGCTTGTCG GACGCAGCAT

```

247

```

51  TGCCGACATC  GCCCAAGATT  TGCCCCGACGG  GGCCGACTGC  CAAATCTGCG
101 CCGTTTACCG  CAACAACCGC  CTCATCGTCC  CCGCGCCGCA  AACCGTCATC
151 ATCGAAGGCG  ACGAAATCCT  ATTTGCCGCC  GCCGCCGAAA  ACATCGGCGC
201 GGTACATACC  GAATTGCGCC  CCAAAGAAAC  CAGCACCCGC  CGCATCATGA
251 TTGCCGGCGG  CGGCAACATC  GGCTACCGTC  TCGCCAAGCA  GTCGAAACAC
301 GCATACAACG  TCAAAATCAT  CGAATGCCGG  CCGCGCCGTG  CCGAATGGAT
351 AGCCGAAAAC  CTGACAACA  CCTCGTCCT  GCAAGGTTCT  GCAACCGACG
401 AAACCCCTGCT  CGACAACGAA  TACATCGACG  AAATCGACGT  ATTCTGCGCC
451 CTGACCAACG  ACGACGAAAG  CAACATTATG  TCCGCCCTTT  TGGCGAAAAA
501 CCTCGGCGCG  AAGCGCGTCA  TCGGCATCGT  CAACCGCTCA  AGCTACGTCG
551 ATTTGCTCGA  AGGCAACAAA  ATCGACATCG  TCGTCTCCCC  CCACCTCATC
601 ACCATCGGCT  CGATACTCGC  CCACATCCGG  CGCGGGGACA  TCGTTGCCGT
651 CCACCCCATC  CGGCGCGGCA  CGGCGGAAGC  CATCGAAGTC  GTCGCACACG
701 GCGACAAAAA  AACTTCCGCC  ATCATCGGCA  GGCGCATCAG  CGGCATCAAA
751 TGGCCCGAAG  GGTGCCACAT  TGCCGCCGTC  GTCCGCGCCG  GAACCGGCGA
801 AACCATTATG  GGACACCATA  CCGAAACCGT  CATCCAAGAC  GGCGACCACA
851 TCATCTTTT  CGTCTCGCG  CGGCGCATCC  TGAACGAACT  GGAAAACTC
901 ATCCAAGTCA  AAATGGGCTT  TTTCGATAA

```

This corresponds to the amino acid sequence <SEQ ID 198; ORF 047.a>:

```

a047.pep
1  MVI IQARRGG LLVGRSIADI AQDLDPGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI GYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDES NIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKK TSA IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRLNLELEKL
301 IQVKMGFFG*

```

m047/a047 96.5% identity over a 312 aa overlap

```

m047.pep      10      20      30      40      50      60
MVI IQARXXGLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIIEGDEILFAA
||||| | ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
a047          10      20      30      40      50      60
MVI IQARRGGLLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIIEGDEILFAA

m047.pep      70      80      90     100     110     120
AAENIGAVIPELRPKETQRNQPXXIMIXGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
||||| | ||||| : || ||||| ||||||| ||||||| ||||||| ||||||| |||||||
a047          70      80      90     100     110
AAENIGAVIPELRPKETSTRR---IMIAGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI

m047.pep     130     140     150     160     170     180
AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDES NIMSALLAKNLGAKRVIGIV
||||| | ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
a047         120     130     140     150     160     170
AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDES NIMSALLAKNLGAKRVIGIV

m047.pep     190     200     210     220     230     240
NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVVAHGDKK
||||| | ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
a047         180     190     200     210     220     230
NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVVAHGDKK

m047.pep     250     260     270     280     290     300
TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRLNEL
||||| | ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
a047         240     250     260     270     280     290
TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRLNEL

m047.pep      310
EKLIQVKMGFFGX
||||| | |||||

```

a047 EKLIQVKMGFFGX  
300 310

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 047 shows 96.2% identity over a 312 aa overlap with a predicted ORF (ORF 047.ng) from *N. gonorrhoeae*:

m047/g045

m047.pep	MVIIQARXXGXLVGRSIADIAQDLPGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA	60
g047	MVIIQARRGGLLVGRSIADIAQDLPGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA	60
m047.pep	AAENIGAVIPELRPKETORNQPPXXIMIXGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI	120
g047	AAENIGAVIPELRPKETSTR---IMTAGGNNICYRLAKQLEHAYNVKIIIECRPRRAEWI	117
m047.pep	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	180
g047	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	177
m047.pep	NRSSVVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	240
g047	NRSSVVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	237
m047.pep	TSIIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRIILNEL	300
g047	TSIIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRIILNEL	297
m047.pep	EKLIQVKMGFFGX	313
g047	EKLIQVKMGFFGX	310

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 199>:

g048.seq

1	ATGCTCGACA AAGGCGAGGA GTTGCCCGTC GATTTCACCA ACCGCCTGAT
51	TTACTACGTC ggcCCcgTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCCG
101	CAGGTCCGAC CACAGCCACC CGCATGGACA AATTTACCCG CCAAATGCTC
151	AAACAAACCG GCCTCTTGGG CATGATCGGC AAATCCGagc gcgGcgcggc
201	cacctGCGAA GCcatCGCCG ACAACAAGGC CGTGACCTC ATGGCAGTCG
251	GCGGCGCGGC ATACCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301	GCGTTCCCCG AATTGGGTAT GGAAGCCGTT TACGAATTG AAGTCAAAGA
351	TATGCCCCGA ACCGTCGCCG TGGACAGCAA AGGCGAATCC ATCCACGCCA
401	CCGCCCCGCG CAAATGGCAG GCGAAAATCG GCATCATCCC CGTCGAGTCT
451	TGA

This corresponds to the amino acid sequence <SEQ ID 200; ORF 048.ng>:

g048.pep

1	MLDKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
51	KQTGLLMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101	AFPELGMEAV YEFVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
151	*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 201>:

m048.seq

1	ATGCTCAACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
51	TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCCG



```
m048.pep
  1  MLNKGEELPV DFTNRLIIFV GPVDPVGDEV VGPAGPTAT RMDKFTRQML
51  EQTDLGLMIG KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 APFELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
151 *
```

```
a048.seq
1  ATGCTCGACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
51  TTACTACGTC GGCCCCG13G ATCCGGTCGG CGACGAAATC GTCGGGCCAG
101 CAGGTCCCGAC CACCGCCACC CGCATGGACA AATTCACCG CCAAATGCTC
151 GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC CGGGCGCGGC
201 CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGTACCTC ATGGCAGTCG
251 GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301 GCGTTCCCGC AATTGGGCAT GGAAGCCATT TACGAATTG AAGTCAAGA
351 CATGCCCGTA ACCGTCGCGC TAGACAGCAA AGGCGAATCC ATCCACGCCA
401 CCGCCCCGCC CCAATGGCAG GCGAAAATCG GCATCATCCC CGTCAAATCT
451 TGA
```

```
a048.pep
  1  MLDKGEELPV  DFTNRLIYYV  GPVDPVGDEI  VGPAGPTTAT  RMDKFTRQML
51  EQTDLLGMIG  KSERGAATCE  AIADNKAVYL  MAVGGAAYLV  AKAIKSSKVL
101 AFPELGMEAI  YEFEVKDMPV  TVAVDSKGES  IHATAPPQWQ  AKIGIIPVKS
151 *
```

	10	20	30	40	50	60
m048.pep	MLNKGEELPVDFTNRLIYYVGPVDPVGDEVVGFAGPTTATRMDKFTROMLEQTDLLGMIG					
a048	MLDKGEELPVDFTNRLIYYVGPVDPVGDEIVGFAGPTTATRMDKFTROMLEQTDLLGMIG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m048.pep	KSERGVATCEAIADNKAVYLMVAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPV					
a048	KSERGAATCEAIADNKAVYLMVAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPV					
	70	80	90	100	110	120
	130	140	150			
m048.pep	TVAVDSKSGESIHATAPRKWQAKIGIIPVESX					
a048	TVAVDSKSGESIHATAPQWQAKIGIIPVKSX					
	130	140	150			

ORF 048 shows 96.4% identity over a 150 aa overlap with a predicted ORF (ORF\_048.ng) from *N. gonorrhoeae*:

m048/q048

	10	20	30	40	50	60
m048.pep	MLNKGEELPVDF	TNRLIYYVGPV	DVPGDEVVGP	PAGPTTATRMD	KFTQRMLEQTD	LLGMIG
g048	MLDKGEELPVDF	TNRLIYYVGPV	DVPGDEVVGP	PAGPTTATRMD	KFTQRMQLKQT	GLLGMIG
	10	20	30	40	50	60
	70	80	90	100	110	120
m048.pep	KSERGVATCEAI	ADNKAVYILMA	VGGAAAYLVAK	AIKSSKVLAF	PELGMEAIYEF	EVKDKMPV
g048	KSERGAATCEAI	ADNKAVYILMA	VGGAAAYLVAK	AIKSSKVLAF	PELGMEAVYEF	EVKDKMPV
	70	80	90	100	110	120
	130	140	150			
m048.pep	TVAVDSKGESI	HATAPRKWQAK	IGIIPVESX			
g048	TVAVDSKGESI	HATAPRKWQAK	IGIIPVESX			
	130	140	150			

q049.seq

1	ATGCGGGGCG	AGGCGTTTGA	TCAACCGTTC	GGTCAGCTCC	TGTTCCGGACA
51	GGCAGAACAC	TTCGCGCCGG	TTGACGGGTT	TCGGGTTAG	GATATTGATT
101	TGGACGGGCA	TCAACGCCTC	TTCCGCACCG	CCTTCGCCCT	TTTCCGCAAC
151	CCCGTCTGCC	GCCGTACCCG	ATTCTGCCCG	ATCGCGGTTT	TCCCGCCCTT
201	CAATCTGTGC	GGTTTCAAAT	TCGCGACTGT	CTTTTTTGCG	ATCGAACCGG
251	ATTCTCCGCC	CGGATTCGAT	GTGTTTTTCC	GAAAccggca	tTTGCAGGGA
301	AGCCTGcgcg	TTGAGCCGAT	TTTCTGAAG	GACGATCATC	GGGTCCGTTT
351	CGACTTCCTC	GCCGCAATCG	GCAACGGCGc	tGTTGTGTTC	TTCTGCCAT
401	TTCTTCGAC	ACGCGTTTAA			

q049.pcp

1 MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRL FRTAFAVFRN  
51 PVCRRTGFCR IGVFPALNLC GFKFGTVFFG IEPDSPPRFD VFFRNRLHQQ  
101 SLRVEPVFLK DDHRVGFDL AAIGNGAVVF FLPLQIRL\*

m049.seq (partial)

1	ATGCGGGCGC	AGGCGTTTGA	TCAGCCGTTC	GGTCAGCTCC	TGTTCCGGACA
51	GGCAGAACAC	TTCGCGCCGG	TTGACGGCTT	TCGGGTTCAG	GATATTGATT
101	TGGACGGGCA	TCAACGTTTC	TTCCGCATCG	TTTTCCCGGT	TTTCCGAAAC
151	CGCCGGCTCA	TTCGTGCCGG	ATTCTGCCTC	GTCCGGGTTT	TCCCGGCTTT
201	CAATCTGTCC	GGTTTTCAAA	TCGACACTGT	CTTTTTTGTT	ATCAAACCGG
251	ATTCTCCGCC	GCGATTCGAT	GTGTTTTTCC	GAAACCGACA	TTTGCAGGGA
301	AGCCTGCGCG	TTGAGCCAGT	TTTCTGAAG	GACGATCATC	GGGTGCGTTT
351	CGACTTCTCT	GCCGCAATCG	GCAACGGCGG	CATTGTGTTC	CTCCTGCCAT
401	TTTTTCACG	AGCCCTT...			

m049.pep (partial)

```

1  MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRF FRIVFPVFRN
51  RRLIRAGFCL VGVFPAFNLS GFKFDTVFFG IKPDSPPRFD VFFRNRLHQQ
101 SLRVEPVFLK DDHRVGFDL AAIGNGGIVF LLPFFOIRL...

```

a049.seq

251

```

      1 ATGCGGGCGC AGGCGTTTGA TCAGCCGTTT GGTGAGCTCC TGTTCGGACA
     51 GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG AATATTGATT
    101 TGGACGGGCA TCAACGCTTC TTCCGCACCG CCTTCGCCGT TTTCCGCAAC
    151 CCGTCTGCC GCCGTACCG ATTCTGCCG ATCGGCGTTT TCCCGCCTT
    201 CAATCTGTCC GGTTCCTCAAT TCGGCACTGT CTTTTTTGGC ATCAAACCGG
    251 ATTCTCCGCC GCGATTCGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
    301 AGCCTGCGCG TTGAGCCAGT TTTCTGAAG GACGATCATC GGGTCGGTTT
    351 CGACTTCCTC GCCGCAATCG GCAACGGCGG CATTGTGTTT CTCCTGCCAT
    401 TTTTTCAGAT ACGCCTT

```

This corresponds to the amino acid sequence <SEQ ID 210: ORF 049.a>:

```

a049.pep
      1 MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ NIDLDGHQRF FRTAFAVERN
     51 PVCRRTRFCR IGVFPAENLS GFKFGTVFFG IKPDSPPRFD VFFRNRLQG
    101 SLRVEPVFLK DHRVGFDFL AAIGNGGIVF LLPFFQIRL

```

m049./a049 90.6% identity over a 139 aa overlap

```

      10      20      30      40      50      60
m049.pep MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLGDGHQRFRRIVFPVFRNRRLIRAGFCL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049      MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLGDGHQRFRRITAFVFRNPVCRTRFCR
      10      20      30      40      50      60

      70      80      90     100     110     120
m049.pep VGVFPAENLSGFKFDTVFFGIKPDSPPRFDVFFRNRLQGSRLRVEPVFLKDDHRVGFDFL
          :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049      IGVFPAENLSGFKFGTVFFGIKPDSPPRFDVFFRNRLQGSRLRVEPVFLKDDHRVGFDFL
      70      80      90     100     110     120

      130     139
m049.pep AAIGNGGIVFLLPFFQIRL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a049      AAIGNGGIVFLLPFFQIRL
      130

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 049 shows 86.3% identity over a 139 aa overlap with a predicted ORF (ORF 049.ng) from *N. gonorrhoeae*:

```

m049/g049
      10      20      30      40      50      60
m049.pep MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLGDGHQRFRRIVFPVFRNRRLIRAGFCL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g049      MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDLGDGHQRLFRITAFVFRNPVCRRTGFCR
      10      20      30      40      50      60

      70      80      90     100     110     120
m049.pep VGVFPAENLSGFKFDTVFFGIKPDSPPRFDVFFRNRLQGSRLRVEPVFLKDDHRVGFDFL
          :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g049      IGVFPALNLCGFKFGTVFFGIEPDSPPRFDVFFRNRLQGSRLRVEPVFLKDDHRVGFDFL
      70      80      90     100     110     120

      130     139
m049.pep AAIGNGGIVFLLPFFQIRL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g049      AAIGNGAVVFLPFLQIRLX
      130     140

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 211>:

```
g050.seq
1  atggggcgCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGg
51  cagccccGAA AAAGccgtgt TGATGGcaaA AGAATCCCTG ATGAGCCACA
101 TCGAcatCca aGaATTGCAG GAAAAAGCCG CGTccggggc ggaattgtcc
151 accaccgaAG ccCTGCGCCT cGAACTCTTT GAAAAGGTCA ACGCGCTGGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAT GCCGCCTCCA AACCGATTGC CATGATTCCC
301 AACTGTGCcg ccacCCGcca cgtcgAATTT GAATTGgACG GCTCAGGtcc
351 TGTCGAactc acgccGCcgc gtgtCGAAGA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 212; ORF 050.ng>:

```
g050.pep
1  MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVED*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 213>:

```
m050.seq
1  ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGTATCG GCATCGGCGG
51  C.agCCgAA AAAGCCGTGC TGATGGCAAA AGAGTCCCTG ATGAGCCACA
101 TCGACATTCA AGAATTGCAG GAAAAGGCCG CGTCCGGCGC GgAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAGTCA ACGCGCTGGG
201 CaTCGGCGCA CAAGGCTTGG GCGGACTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA TCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351 TGTCGAACTC ACGCCGCCGC GCGTCGAAGA TGGCCCGATT TGA
```

This corresponds to the amino acid sequence <SEQ ID 214; ORF 050>:

```
m050.pep
1  MGAGWCPPGI LGIGIGGXAE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDGPI *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 215>:

```
a050.seq
1  ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGG
51  TACGCCCGAA AAAGCCGTGT TGATGGCGAA AGAATCCCTG ATGAGCCACA
101 TCGACATCCA AGAATTGCAG GAAAAGCCG CGTCCGGCGC GGAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAGTCA ACGCGCTAGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351 TGTCGAACTC ACGCCGCCGC GCGTCGAAGA CTGGCCC
```

This corresponds to the amino acid sequence <SEQ ID 216; ORF 050.a>:

```
a050.pep
1  MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDWP
```

m050/a050 97.7% identity over a 129 aa overlap

```

          10      20      30      40      50      60
m050.pep  MGAGWCPPGILGIGIGGXAEKAVLMAKESLSMSHIDIQELQEKAASGAELSTTEALRLELF
          |||||
a050       MGAGWCPPGILGIGIGGTPEKAVLMAKESLSMSHIDIQELQEKAASGAELSTTEALRLELF
```

253

	10	20	30	40	50	60
	70	80	90	100	110	120
m050.pep	EKNVALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPV					
a050	EKNVALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPV					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
a050	TPPRVEDWP					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 050 shows 98.4% identity over a 127 aa overlap with a predicted ORF (ORF 050.ng) from *N. gonorrhoeae*:

m050/g050

	10	20	30	40	50	60
m050.pep	MGAGWCPPGILGIGIGGXAEKAVLMAKESLSMHSIDIQELQEKAASGAELSTTEALRL					
g050	MGAGWCPPGILGIGIGGTPEKAVLMAKESLSMHSIDIQELQEKAASGAELSTTEALRL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m050.pep	EKNVALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPV					
g050	EKNVALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPV					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
g050	TPPRVEDX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 217>:

g050-1.seq

```

1  ATGACCGTTA TCAAGCAAGA AGACTTTATT CAAAGTATCT GCGATGCGCTT
51  CCAATTCATC AGCTACTACC ATCCAAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCCGCCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCCGAA AACACCGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG GATGTGCAAT
251 GGGATGCGGA CATGAGCGTG GAAAAGATGG TTAACGAAGG CGTACGCCGC
301 GCCTACACTT GGAAGGCAA CACCCTGCGC GCTTCGTCC TCGCCGATCC
351 GCGCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCGCC GTCCATCCAA
401 TGAGCATCGT GCGGCGCGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GCGGCTCTG AAAACAAATC CAAACTCGCT ATGCTCAACC CTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACGATGGGC GCGGGCTGGT
551 GTCTCTCCCG CATCTTGGGC ATCGGCATCG GCGGCACgCC CGAAAAAGCC
601 GTGTTGATGG cgaAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAGAG GTCAACGCGC TGGGCATCGG CGCGCAAGGC
751 TTGGGCGGTC TGACCAACCG GTTGGACGTG AAAATCCTCG ATTACCCGAC
801 CCATGCCGCC TCCAAACCGA TTGCCATGAT TCCCAACTGT GCCGCCACCC
851 GCCACGTCGA ATTTGAATTG GACGGCTCAG GTCTGTGCGA ACTCAGCCCG
901 CCGCGCGTCG AAGACTGACC CGATCTGACT TACAGCCCGC ACAACGGCAA
951 ACGCGTCGAT GTCGATAAGC TGACCAAGA AGAAGTGCCA AGCTGGAAAA
1001 CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCCG CCGCGATGCC
1051 GCGCACAAAC GCCTCGTCAA TATGCTCGAC AAAGCGGAGG AGTTGCCCGT
1101 CGATTTACAC AACCGCCTGA TTTACTACGT CGGCCCGGTC GATCCGGTCG
1151 GCGATGAAGT CGTCGGTCCC GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTTACCC GCCAAATGCT CAAACAAACC GGCCTCTTGG GCATGATCGC

```

```

1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAAACAAGG
1301 CCGTGTACCT CATGGCAGTC GCGGCGCGG CATACCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGGTTCCCG GAATTGGGTA TGGAAAGCCGT
1401 TTACGAATTT GAAGTCAAAG ATATGCCCGT AACCGTCGCC GTGGACAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCCCG GCAATGGCA GGCAGAAATC
1501 GGCATCATCC CCGTCGAGTC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 218; ORF 050-1.ng>:

**g050-1.pep**

```

1  MTVIKQEDFI QSIDAFQFI SYHHPKDYID ALYKAWQKEE NPAAKDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVG M DVQWDADMSV EKMVNEGVRR
101 AYTWEGNTLR ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTC AARG
151 GGSSENKSLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELF EK VNALGIGAOG
251 LGGTTTVLDV KILDYPTHA A SKPIAMIPNC AATRHVEFEL DSGSPVELTP
301 PRVED*PDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351 AHKRLVNMLD KGEELPVDF T NRLIYYVGPV DPVGDEVVGP AGPTTATRMD
401 KFTRQMLKQT GLLGMIGKSE RGAATCEAIA DNKAVYLM AV GGAAYLVAKA
451 IKSSKVLAFP ELGMEAVYEF EVKDMPTVTA VDSKGESIHA TAPRKWQAKI
501 GIIPVES*

```

**g050-1/p14407**

```

sp|P14407|FUMB_ECOLI FUMARATE HYDRATASE CLASS I, ANAEROBIC (FUMARASE)
>gi|280063|pir|B44511 fumarate hydratase (EC 4.2.1.2) fumB, iron-dependent - Escherichia coli
>gi|146048 (M27058) anaerobic class I fumarase (EC 4.2.1.2) [Escherichia coli] Length = 548
Score = 172 bits (432), Expect = 4e-42
Identities = 138/488 (28%), Positives = 216/488 (43%), Gaps = 22/488 (4%)

```

```

Query: 11 QSIDAFQFISYHHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAENHRPICQDTG 70
      Q+ DA + H K L+ E + K Q L NS + A+ P CQDTG
Sbjct: 53 QAFHDASFMLRPAHQKQVAAAILHDPEASEND---KYVALQLRNSEIAAKGVLPCTQDTG 109

```

```

Query: 71 IATVFLKVGMDVQWDADMSVEKMVNEGVRRAYTWEGNTLRASVLADPAGKRONTKDNTPA 130
      A + K G V W E+ +++GV Y E N + A K NT N PA
Sbjct: 110 TAIIVGKKGQRV-WTGGGD-EETLSKGVYNTYI-EDNLRYSQNAALDMYKEVNTGTNLPA 166

```

```

Query: 131 VIHMSIVPGGKVEVTC AAKGGGSENKSL-----AMNPSDNIVDWVLKTIPTMGAGWCP 185
      I + V G + + C AKGGGS NK+ L A+L P + +++++ + T+G CP
Sbjct: 167 QIDLYAVDGDEYKFLCVAKGGGSANKTYLYOETKALLTPG-KLKNFLVEKMRTLGTAAACP 225

```

```

Query: 186 PXXXXXXXXXTPEKAVLMAKESLSHSHIDIQELQEKAAAGELSTTEALRLELFKVNXXX 245
      P T + L + +H EL + + L EL E+
Sbjct: 226 PYHIAFVIGGTS AETNLKTVKLASAHY-YDELPTEGNEHQAFRDVQLEQELLEEAQKLG 284

```

```

Query: 246 XXXXXXXXTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSG---PVELTPP 301
      D++++ P H AS P+ M +C+A R+++ +++ G +E P
Sbjct: 285 LGAQFGGKYFAH-DIRVIRLPRHGASCPVGMGVSCSADRNIKAKINREGIWKLEHNPG 343

```

```

Query: 302 RVEDXPDLTYSPDNGKRVDVDKLTKE---EVASWKTGDVLLNGKILTGRDAAHKRLVNM 358
      + +VD+++ KE +++ + L L G I+ GRD AH +L +
Sbjct: 344 QYIPQELRQAGEGEAVKVDLNRPMKEILAQLSQYPVSTRLSLTGTIIVGRDIAHAKLKL 403

```

```

Query: 359 LDKGEELPVDFNRLIYYXXXXXXXXXXXXXXXXTTATRMDKFTTRQMLKQTGLLGMIGK 418
      +D G+ELP + IYY TTA RMD + + G + M+ K
Sbjct: 404 IDAGKELPOYIKDHP IYYAGPAKTPAGYPSGSLGPTTAGRMDSYVDLLQSHGGSMIMLAK 463

```

```

Query: 419 SERGAATCEAIADNKAVYLMVAVGG-AAYLVAKAIKSSKVLAFPPELGMEAVYEFVVDMPV 477
      R +A + YL ++GG AA L ++IK + +A+PELGMEA+++ EV+D P
Sbjct: 464 GNRSQQVTDACHKHGGFYLSIGGPAAVLAQQSIKHLECVAYPELGMEAIWKIEVEDFPA 523

```

```

Query: 478 TVAVDSKG 485
      + VD KG
Sbjct: 524 FILVDDKG 531

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 219>:

**m050-1.seq**

```

1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
51  CCAATTCATC AGCTACTATC ATCCCAAAGA CTACATCGAC GCGTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCTGCCG CCAAAGACGC GATACGCAG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCGGAA AACACCGGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG AACGTCCAAT
251 GGGATGCGGA CATGAGCGTG GAAGAGATGG TTAACGAAGG CGTACGCCGC

```

```

301 GCCTACACTT GGAAGGCAA TACGCTGCGC GCTCCGTC TCGCCGATCC
351 GGCCGGGAAA CGCCAAAACA CCAAAGACAA CACCCCGGCC GTCATCCATA
401 TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGCTCTG AAAACAAATC CAAACTCGCC ATGCTCAATC CTTCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACCATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGT ATCGGCATCG GCGGCACGCC CGAAAAGGCC
601 GTGCTGATGG CAAAGAGTC CCTGATGAGC CACATCGACA TTCAAGAATT
651 GCAGGAAAAG GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAAAA GTCAACGCGC TGGGCATCGG CGACAAGGC
751 TTGGGCGGAC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTATCCGAC
801 CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAAGTGC GCGGCCACCC
851 GCCACGTCGA ATTGAATTG GACGGCTCAG GCCCTGTCGA ACTCAGCCCG
901 CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCGG ACAACGTTTA
951 ACGCGTCGAT GTGACAAGC TGACCAAGA AGAAGTGGCA AGCTGAAAA
1001 CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACC GGCGATGCC
1051 GCACACAAAC GCCTCGTCGA TATGCTCAAC AAAGGCGAAG AATTGCCCGT
1101 CGATTTCACC AACCGCCTGA TTTACTACGT CCGCCCGGTC GATCCGGTCG
1151 CGCATGAAGT CGTCGGTCCG GCAGGTCCGA CCACAGCCAC CGCATGGAC
1201 AAATTCACCC GCCAAATGCT CGAACAAAC GACTCTTGG GCATGATCGG
1251 CAAATCCGAG CCGGCGTGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301 CCGTGTAACC CATGGCAGT GCGGCGCGG CGTATCTCGT GGCAAAAGCC
1351 ATCAAACTTT CCAAAGTCTT GCGGTTCGCC GAATTGGGCA TGGAGCCAT
1401 TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGATAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGG CAAATGGCA GCGAAAATC
1501 GGCATCATCC CCGTCGAATC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 220; ORF 050-1>:

m050-1.pep

```

1  MTVIKQEDFI QSIDAFQFI SYHHPKDYID ALYKAWQKEE NPAARDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVG M NVQWDADMSV EEMVNEGVRR
101 AYTWEGNTRL ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTCAAKG
151 GGSSENKSLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELOEK AASGAELSTT EALRLELFEX VNALGIGAQQ
251 LGGLTTVLVDV KILDYPTHA A SKPIAMIPNC AATRHVEFEL DSGSPVELTP
301 PRVEDWPDLT YSPDNKRV D VDKLTKEEVA SWKTGDVLL NGKILTGRDA
351 AHKRLVMDLN KGEELPVDF N RLIIYVGPV DPVGDVEVGP AGPTTATRMD
401 KFTAMLEQT DLLGKIKSE RGVATCEAIA DNKAVYLMV GGAAYLVAKA
451 IKSSKVLAFP ELGMEAIYEF EVKDMPTVA VDSKGESIHA TAPRKWQAKI
501 GIIPVES*

```

m050-1/g050-1 98.2% identity in 507 aa overlap

	10	20	30	40	50	60
m050-1.pep	MTVIKQEDFIQSIDAFQFISYHHPKDYIDALYKAWQKEENPAARDAMTQILVNSRMCAE					
g050-1	MTVIKQEDFIQSIDAFQFISYHHPKDYIDALYKAWQKEENPAARDAMTQILVNSRMCAE					
	10	20	30	40	50	60
m050-1.pep	NNRPICQDTGIATVFLKVG MNVQWDADMSVEEMVNEGVRRAYTWEGNTRLASVLADPAGK					
g050-1	NNRPICQDTGIATVFLKVG MDVQWDADMSVEKMNVEGVRRAYTWEGNTRLASVLADPAGK					
	70	80	90	100	110	120
m050-1.pep	NNRPICQDTGIATVFLKVG MNVQWDADMSVEEMVNEGVRRAYTWEGNTRLASVLADPAGK					
g050-1	NNRPICQDTGIATVFLKVG MDVQWDADMSVEKMNVEGVRRAYTWEGNTRLASVLADPAGK					
	70	80	90	100	110	120
m050-1.pep	RQNTKDNTPAVIHMSIVPGGKVEVTC AAKGGGSENKSLAMLNPSDNIVDWVLKTIPTMG					
g050-1	RQNTKDNTPAVIHMSIVPGGKVEVTC AAKGGGSENKSLAMLNPSDNIVDWVLKTIPTMG					
	130	140	150	160	170	180
m050-1.pep	RQNTKDNTPAVIHMSIVPGGKVEVTC AAKGGGSENKSLAMLNPSDNIVDWVLKTIPTMG					
g050-1	RQNTKDNTPAVIHMSIVPGGKVEVTC AAKGGGSENKSLAMLNPSDNIVDWVLKTIPTMG					
	130	140	150	160	170	180
m050-1.pep	AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELOEKAASGAELSTTEALRLELFEX					
g050-1	AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELOEKAASGAELSTTEALRLELFEX					
	190	200	210	220	230	240
m050-1.pep	AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELOEKAASGAELSTTEALRLELFEX					
g050-1	AGWCPPGILGIGIGGTPEKAVLMAKESLSHIDIQELOEKAASGAELSTTEALRLELFEX					
	190	200	210	220	230	240
m050-1.pep	VNALGIGAQGLGGLTTVLVDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGS GPVELTP					
g050-1	VNALGIGAQGLGGLTTVLVDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGS GPVELTP					
	250	260	270	280	290	300
m050-1.pep	VNALGIGAQGLGGLTTVLVDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGS GPVELTP					
g050-1	VNALGIGAQGLGGLTTVLVDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGS GPVELTP					
	250	260	270	280	290	300

256

	310	320	330	340	350	360
m050-1.pep	PRVEDWPDLT	YSPDNGKRV	VDKLTKEEVAS	WKTGDVLL	LLNGKIL	TGRDAAHKRLV
g050-1	PRVEDXPDLT	YSPDNGKRV	VDKLTKEEVAS	WKTGDVLL	LLNGKIL	TGRDAAHKRLV
	310	320	330	340	350	360
	370	380	390	400	410	420
m050-1.pep	KGEELPVDFT	NRLIYYVGP	VPDVPVGD	EVVGPAGP	TTATRMDF	TRQMLEQTD
g050-1	KGEELPVDFT	NRLIYYVGP	VPDVPVGD	EVVGPAGP	TTATRMDF	TRQMLEQTD
	370	380	390	400	410	420
	430	440	450	460	470	480
m050-1.pep	RGVATCEA	IADNKAVY	LMVGGAA	YLVAKAIK	SSKVLAF	PELGMEAI
g050-1	RGVATCEA	IADNKAVY	LMVGGAA	YLVAKAIK	SSKVLAF	PELGMEAI
	430	440	450	460	470	480
	490	500				
m050-1.pep	VDSKGESI	HATAPRK	WQAKIGI	IPVESX		
g050-1	VDSKGESI	HATAPRK	WQAKIGI	IPVESX		
	490	500				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 221>:

a050-1.seq

1	ATGACCGTCA	TCAAACAGGA	AGACTTTATC	CAAAGCATTT	CGGATGCCTT
51	CCAATTTCATC	AGCTACTACC	ATCCCAAAGA	CTACATCGAC	CGGCTTTATA
101	AGGCGTGGCA	GAAGGAAGAA	AACCCCGCCG	CCAAAGACGC	GATGACGCAG
151	ATTTTGGTCA	ACAGCCGCAT	GTGTGCCGAA	AACAACCGCC	CCATCTGCCA
201	AGATACCGGT	ATCGCGACCG	TGTTTTTGAA	AGTCGGTATG	GATGTGCAAT
251	GGGATGCAGA	CATGAGCGTC	GAAGAGATGG	TTAACGAAGG	CGTGCGCCGC
301	GCCTACACTT	GGGAAGGCAA	TACGCTGCGC	GCTTCCGTTT	TCGCCGACCC
351	CGCCGGCAAA	CGCCAAAATA	CCAAAGACAA	CACGCCCGCC	GTCATCCATA
401	TGAGCATCGT	GCCGGGCGAC	AAAGTCGAAG	TAACCTGCGC	GGCAAAAGGC
451	GGCGGTTCCTG	AAAACAAATC	CAAACTCGCC	ATGCTCAACC	CTTCCGACAA
501	CATCGTTCGAT	TGGGTATTGA	AAACCATTC	GACCATGGGC	CGCGGCTGGT
551	GTCCCTCCCG	CATCTTGGGC	ATCGGCATCG	GCGGTACGCC	CGAAAAAGCC
601	GTGTTGATGG	CGAAAGAATC	CCTGATGAGC	CACATCGACA	TCCAAGAATT
651	GCAGGAAAAA	GCCCGCTCCG	GCGCGGAATT	GTCCACCACC	GAAGCCCTGC
701	GGCTCGAACT	CTTTGAAAAA	GTCAACGCGC	TAGGCATCGG	CGCGCAAGGC
751	TTGGGCGGTC	TGACCACCGT	GTTGGACGTG	AAAATCCTCG	ATTACCCGAC
801	CCACGCCGCC	TCCAAACCGA	TTGCCATGAT	TCCGAAGTGC	GCCGCCACCC
851	TGCACGTCGA	ATTTGAATTG	GACGGCTCAG	GCCCTGTCGA	ACTCACGCCG
901	CCGCGCGTCG	AAGACTGSCC	CGATTTGACT	TACAGCCCCG	ACAACGGCAA
951	ACGCGTCGAT	GTCGACAAGC	TGACCAAAGA	AGAAGTGCCA	AGCTGGAAAA
1001	CCGGCGACGT	ATTGCTGTTG	AACGGCAAAA	TCCTCACCGG	CCGCGATGCC
1051	GCACACAAAC	GCCTCGTCGA	TATGCTCGAC	AAAGGCGAAG	AATTGCCCGT
1101	CGATTTTCACC	AACCGCCTGA	TTTACTACGT	CGGCCCGGTC	GATCCGGTCG
1151	GGGACGAAAT	CGTCGGCCCA	GCAGGTCCGA	CCACCGCCAC	CCGCATGGAC
1201	AAATTCACCC	GCCAAATGCT	CGAACAAACC	GACCTCTTGG	GCATGATCGG
1251	CAAATCCGAG	CGCGGCGCGG	CCACCTGCGA	AGCCATCGCC	GACAACAAAG
1301	CCGTGTACCT	CATGGCAGTC	GCGGGCGCGG	CGTATCTCGT	GGCAAAAGCC
1351	ATCAAATCTT	CCAAAGTCTT	GGCGTTCCCC	GAATTGGGCA	TGGAAGCCAT
1401	TTACGAATTT	GAAGTCAAAG	ACATGCCCGT	AACCGTCGCC	GTAGACAGCA
1451	AAGGCGAATC	CATCCACGCC	ACCGCCCCGC	CCCAATGGCA	GGCGAAAAATC
1501	GGCATCATCC	CCGTCAAATC	TTGA		

This corresponds to the amino acid sequence <SEQ ID 222; ORF 050-1.a>:

a050-1.pep

1	MTVIKQEDFI	QSICDAFQFI	SYHPKDYID	ALYKAWQKEE	NPAAKDAMTQ
51	ILVNSRMCAE	NNRPICQDTG	IATVFLKVG	MDVQWADMSV	EEMVNEGVRR
101	AYTWEGNTLR	ASVLADPAGK	RQNTKDNTPA	VIHMSIVPGD	KVEVTCAAKG
151	GSGENKSLA	MLNPSDNIVD	WVLKTIPTMG	AGWCPPGILG	IGIGGTPEKA
201	VLMAKESLMS	HIDIQELQEK	AASGAELSTT	EALRLELFKE	VNALGIGAQG
251	LGGLTTVLDV	KILDYPHTAA	SKPIAMIPNC	AATRHVEFEL	DGSGPVELTP
301	PRVEDWPDLT	YSPDNGKRV	VDKLTKEEVA	SWKTGDVLL	NGKILTRDA
351	AHKRLVDMLD	KGEELPVDFT	NRLIYYVGP	DPVGDIVGP	AGPTTATRMD
401	KFTRQMLEQT	DLGGMIGKSE	RGAATCEAIA	DNKAVYLMV	GGAAAYLVAKA
451	IKSSKVLAF	ELGMEAIYEF	EVKDMPVTVA	VDSKGESIHA	TAPPOWQAKI
501	GIIPVKS*				



257

a050-1/m050-1 98.4% identity in 507 aa overlap

```

      10      20      30      40      50      60
a050-1.pep MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
|||||
m050-1      MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
      10      20      30      40      50      60

      70      80      90     100     110     120
a050-1.pep NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
|||||
m050-1      NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
      70      80      90     100     110     120

      130     140     150     160     170     180
a050-1.pep RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWLKTIPMTG
|||||
m050-1      RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWLKTIPMTG
      130     140     150     160     170     180

      190     200     210     220     230     240
a050-1.pep AGWCPPGILGIGGTPPEKAVLMAKESLMASHIDIQELQEKAASGAELSTTEALRLELFEK
|||||
m050-1      AGWCPPGILGIGGTPPEKAVLMAKESLMASHIDIQELQEKAASGAELSTTEALRLELFEK
      190     200     210     220     230     240

      250     260     270     280     290     300
a050-1.pep VNALGIGAQGLGGLTTVLDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVELTP
|||||
m050-1      VNALGIGAQGLGGLTTVLDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGSGPVELTP
      250     260     270     280     290     300

      310     320     330     340     350     360
a050-1.pep PRVEDWPDLTYSPPNGKRVVDKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLE
|||||
m050-1      PRVEDWPDLTYSPPNGKRVVDKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLE
      310     320     330     340     350     360

      370     380     390     400     410     420
a050-1.pep KGEELPVDFTNRLIYYVGPVDPVGDVGVGAGPTTATRMDFTRQMLEQTDLLGMIGKSE
|||||
m050-1      KGEELPVDFTNRLIYYVGPVDPVGDVGVGAGPTTATRMDFTRQMLEQTDLLGMIGKSE
      370     380     390     400     410     420

      430     440     450     460     470     480
a050-1.pep RGAATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVKDMPVTVA
||:|||||
m050-1      RGVATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVKDMPVTVA
      430     440     450     460     470     480

      490     500
a050-1.pep VDSKGESIHATAPPQWQAKIGIIPVKSX
|||||
m050-1      VDSKGESIHATAPRKWQAKIGIIPVESX
      490     500

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 223>:

g052.seq

```

1  ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
51  CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
151 AAGGGGTTGG ACGGCGTTTC CAAAACAGC AGTTTGGTGT TGGCTTTGAC
201 GGCGGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CTCACTCCGA
251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
351 CGCCGCCTGA

```

This corresponds to the amino acid sequence &lt;SEQ ID 224; ORF 052.ng&gt;:

g052.pep

```

1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
101 RLRLETTWSP ACRKVKNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 225>:

```

m052.seq
1 ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
51 CTGAGAGCCG ACAGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
151 AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201 GGCGGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CTCACTCCGA
251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
351 CGCCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 226; ORF 052>:

```

m052.pep
1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
101 RLRLETTWSP ACRKVKNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 227>:

```

a052.seq
1 ATGGCTTTGG TCGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG
51 CTGAGAGCCG ACAGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCCTCC
151 AAGGGATTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201 GGCGGCTTTC CATTCTTTA TATCAGTCGG CGACACGTGA CTCACTTCGA
251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301 AGGCTGCGGC TGGAAATCAC ATGGTCGCCC GCCTGCAAAA AGGTGAAAAA
351 CGCCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 228; ORF 052.a>:

```

a052.pep
1 MALVAEETEI SAPCFKG*EP TGDSRLLSTT KSAPMPCANS AKASKSATSP
51 KGLDGVSKNS SLVLALTAAF HSFISVGD* LTSMNPVLVTM LLIKPTVVPN
101 RLRLEITWSP ACKKVKNA*

```

m052/a052 95.8% identity over a 119 aa overlap

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
a052	MALVAEETEI SAPCFKGXEPGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m052.pep	SLVLALTAAFHSFISVGDTRLTPMPNLVTMLLIKPTVVPNRLRLLETTWSPACRKVKNAAX					
a052	SLVLALTAAFHSFISVGDXTLTSMPNLVTMLLIKPTVVPNRLRLLEITWSPACKKVKNAAX					
	70	80	90	100	110	120

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 052 shows 95.8% identity over a 119 aa overlap with a predicted ORF (ORF 052.ng) from *N. gonorrhoeae*:

m052/g052

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					

259

```

g052      MALVAEETEISAPCFKGCEPTGDSRLSTTKSAPMPCANSASAKSATS PKGLDGVSKNS
              10      20      30      40      50      60

              70      80      90      100     110     120
m052.pep   SLVLALTAAFHSFISVGDTWLTSMPLATMLLIKPTVVPNRLRLEITWSPACKXVKNAAAX
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g052      SLVLALTAAFHSFISVGDTRLTPMPLVMTMLLIKPTVVPNRLRLETTWSPACRKVNAAAX
              70      80      90      100     110     120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 229>:

```

g073.seq
1   ATGTGTATGC CATACGCAAT AAGGGTTTCA GACGGCATCT GCCGCATTTT
51  TCCGCCGATG CCGTCTGAAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
151 AGTCCGGGGC GGATCGGGC GCGGAGTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGTT GTCCACGGAT TGGTGATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAAAAGAGT TCCACGACTT TCCACGCGGC TGCTTGCTCG
301 GCGACTTCAA AACCGATGAC GATGCCGCCG CCGTTTGTCT GTTTGCGGAT
351 AAGCTCCGCC TGCGGATGGT CGGGCAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 230; ORF 073.ng>:

```

g073.pep
1   MCMPIAIRVS DGICRIFPPM PSETRNQRAS ACFKSSIKSP TYSKPTDRRT
51  SPGRIPAAFS SSGCILPCVV VHGLVMVERT SPRLAVREKS STTFHAAAWS
101 ATSKPMTMPP PFCCLRISSA CGWSGNPV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 231>:

```

m073.seq
1   ATGTGTATGC CATATAAGAT AAGGGTTTCA GACGGCATCT GCTGTCCAAT
51  GCGGTCTGAA ACACGCAATC AGCGTGCGAG TGCTGTGTTT AAATCGTCAA
101 TCAAATCGCC AACATATTCC AAACCGACCG ACAGGCGCAC CAATCCGGGG
151 CGGATGTTGG CGGCGAGTTT TTCTTCGGGC TGCATCCTGC CGTGCGTGTT
201 TGTCCACGGG TGGTAATGG TCGAGCGCAC GTCACCGAGG TTGGCGGTGC
251 GGGAAAAGAG TTCCACGCCG TCCACAACCT TCCACGCCGC TTCTTGATCG
301 GCAACTTCAA AGCCGATGAC GATGCCGCCG CCGTTTGTCT GTTTGCGGAT
351 AAGCGCCGCC TGAGGATGGT CGGACAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 232; ORF 073>:

```

m073.pep
1   MCMPIKIRVS DGICCPMPSE TRNQRASACF KSSIKSPTYS KPTDRRTNPG
51  RMLAASFSSG CILPCVVVHG WVMVERTSPR LAVREKSSTP STTFHAASXS
101 ATSKPMTMPP PFCCLRISAA XGWSNDPV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 233>:

```

a073.seq
1   ACGTGTATGT CATATAAGAT AAGGGTTTCA GACGGCATTT GCGGTGTTTT
51  TCCGCCGATG CCGTCTGAA. CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
151 AATCCGGGGC GGATGTTGGC GCGGAGTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGTT GTCCACGGAT GGGTAATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAGAAAAGT TCGACGCCGT CCACGACTTT CCACGCGGCT
301 GCTTGGTGG CGACTTCAA GCGGATGACG ATGCCGCCGC CGTTTGTCTG
351 TTTGCGGATA AGCTCCGCCT GAGGATGGTC GGGTAATCCG GTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 234; ORF 073.a>:

```

a073.pep
1   TCMSYKIRVS DGICGVFPPM PSEXNRQRAS ACFKSSIKSP TYSKPTDRRT
51  NPGRMLAASF SSGCILPCVV VHGWVMVERT SPRLAVREKS STPSTTFHAA
101 AWSATSKPMT MPPFCLRLI SSA*GWSGNP V*

```

m073/a073 92.3% identity over a 130 aa overlap

260

```

      10      20      30      40      50
m073.pep  MCMPIKIRVSDGICCC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
           || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a073      TCMSYKIRVSDGICGVFPMPSEXRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
           10      20      30      40      50      60

      60      70      80      90      100     110
m073.pep  SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCRLI
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a073      SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAAWSATSKPMTMPPPFCCRLI
           70      80      90      100     110     120

      120     129
m073.pep  SAAXGWSNDNPVX
           |: |||| ||||
a073      SSAXGWSGNPVX
           130

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 073 shows 87.0% identity over a 131 aa overlap with a predicted ORF (ORF 073.ng) from *N. gonorrhoeae*:

m073/g073

```

      10      20      30      40      50
m073.pep  MCMPIKIRVSDGICCC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g073      MCMPIAIRVSDGICRIFFPMPSETRNQRASACFKSSIKSPTYSKPTDRRTSPGRIPAAASF
           10      20      30      40      50      60

      60      70      80      90      100     110
m073.pep  SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCRLI
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g073      SSGCILPCVVVHGLVMVERTSPRLAVREKSST---TFHAAWSATSKPMTMPPPFCCRLI
           70      80      90      100     110

      120     129
m073.pep  SAAXGWSNDNPVX
           |: |||| ||||
g073      SSACGWSGNPVX
           120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 235>:

g075.seq

```

1  ATGCCGCCTT ACTTCATCAC CCTCTTAACG ATGGAAAATA CAAAAAGCGC
51  GCGGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCCGTTTCGG GCAACGCTGC GTTTGCCTGT
151 GCGGCCAAAG CCAGCGGGGC GGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTCATGAT ATTTTCCTTT ACGAAATTT
251 TAAAAAATG TGTTGCGGG CTTGTGAAG GTTTAGAGA CCGCTGCCG
301 GGCCTCTTAA ACTTAATCTT CTTTTTCGTA GAATCCGAAA ATTACAAATT
351 CCCCCCTAT CTCTTCCAAT GCCGAGCTAA AAGCGTCTTC ATAGCTGTCA
401 TATTACCGG CTGA

```

This corresponds to the amino acid sequence <SEQ ID 236; ORF 075.ng>:

g075.pep

```

1  MPPYFITLET MENTKSAKT PTTIQASIP SAFAASKAFF AVSGNAAFAC
51  AAKASGAAVT TASFAPYLRQ VLINFMIFSF TKFLKKVCVG LCEGFRDRLP
101 GLNLNLIFFV ESENYKFPAY LFQCRKSVF IAVIFTG*

```

This corresponds to the amino acid sequence <SEQ ID 238; ORF 075>:

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 075 shows 65.7% identity over a 137 aa overlap with a predicted ORF (ORF 075.ng) from *N. gonorrhoeae*:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 239>:

This corresponds to the amino acid sequence <SEQ ID 240; ORF 075.a>:

a075.pep

```
1  MPSYFITLLT MEKTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNVAFAC
51  AAKARGAAVT TASFAPYLQ  VLINEMIFSF KKCLAVMDGA FFRRPPNIRK
101 SVFOKSEYDK FVLVADFFQT CVNRRFEVVE IIGIGD*
```

m075/a075 98.5% identity over a 136 aa overlap

```

              10      20      30      40      50      60
m075.pep    MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
|||||
a075        MPSYFITLLTMEFKTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
              10      20      30      40      50      60

              70      80      90     100     110     120
m075.pep    TASFAPYLRLQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
|||||
a075        TASFAPYLRLQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
              70      80      90     100     110     120

              130
m075.pep    CVNRFFEVEIIGIGDX
|||||
a075        CVNRFFEVEIIGIGDX
              130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 241>:

```

g080.seq
1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTTCG TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAAGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA TACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGTTGCA CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAA TGTCGTTTGG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACgAGa TGAAACGCCT CCgGCTTTTT
601 ACcgAAGCGT GGCAGCATCT gttgcGTAAG AATAAAATC GGTtATCCTA
651 TGTGGATATG Aggtataagg acggatttTC agtcccccat gctCCCGACG
701 GTTTACCCGA AAAAGAAATcc gAAGAAtatt gggaaacagg ttgggacata
751 ttacggcctg gcgtcggaaa cggttcgcag caaatttcaa tcagttatAA
801 GGGCAGacga acaatggaac AGcagtaa

```

This corresponds to the amino acid sequence <SEQ ID 242; ORF 080.ng>:

```

g080.pep
1  MWDNAEAMER LTRWLLVMMa MLLAASGLVW FYNsNHLpVK QVSLKGNLVY
51  SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYpWIASVM VRRRFpDTVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RpgMPVFRGA EGtSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWNVVL DNGITVRLGR ENEMKRLRLf
201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVPH APDGLPEKES EEYWEQVWDI
251 LRPgVGNGST QISISYKGRR TMEQQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 243>:

```

m080.seq
1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTTCG TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGACATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGTTGCA CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTGG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT

```

601 ACCGAAGCGT GGCAGCATCT GTTGCCTAAA AATAAAAAATC GGTTATCCTA  
 651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTTCCGACG  
 701 GTTTACCCGA AAAAGAATCC GAAGAATAG

This corresponds to the amino acid sequence <SEQ ID 2441; ORF 080>:

m080.pep  
 1 MWDNAEAMER LTRWLLVMM MLLAASGLVW FYNNSHLPVK QVSLKGNLVY  
 51 SDKKTLGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDVE  
 101 VVLTERRKPA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR  
 151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF  
 201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY ASDGLPEKES EE\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 080 shows 97.9% identity over a 242 aa overlap with a predicted ORF (ORF 080.ng)

from *N. gonorrhoeae*:

m080/g080

m080.pep	10	20	30	40	50	60
	MWDNAEAMERL	TRWLLVMMAM	LLAASGLVWF	YNSHLPVKQV	SLKGNLVYSD	KKTLGSLA
080	MWDNAEAMERL	TRWLLVMMAM	LLAASGLVWF	YNSHLPVKQV	SLKGNLVYSD	KKALGSLA
	10	20	30	40	50	60
m080.pep	70	80	90	100	110	120
	KEYIHGNILRT	DINGAQEAYR	RYPWIASVMV	RRRFPDVEV	VLTERRKPA	RWGDHALVDG
080	KEYIHGNILRT	DINGAQEAYR	RYPWIASVMV	RRRFPDVEV	VLTERRKPA	RWGDHALVDG
	70	80	90	100	110	120
m080.pep	130	140	150	160	170	180
	EGNVFEARLD	RPGMPVFRGA	EGTSAEMLRR	YDEFSTVLAK	QGLGIKEMTY	TARSAWIVVL
080	EGNVFEARLD	RPGMPVFRGA	EGTSAEMLRR	YDEFSTVLAK	QGLGIKEMTY	TARSANNVVL
	130	140	150	160	170	180
m080.pep	190	200	210	220	230	240
	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLRK	NKNRLSYVDM	RYKDGFSVRY	ASDGLPEKES
080	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLRK	NKNRLSYVDM	RYKDGFSVRY	PHAPDGLPEKES
	190	200	210	220	230	240
m080.pep	EEX					
080	EEYWEQVWDIL	RPGVNGSTQI	SISYKGRRTM	EQQX		
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 245>:

a080.seq  
 1 ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT  
 51 CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGTTTGG TTTTACAATT  
 101 CGAATCATCT GCCCGTCAAG CAGGTGTGCG TGAAGGGCAA CCTAGTTTAT  
 151 TCCGATAAGA AAGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA  
 201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCTTAC CGCCGGTATC  
 251 CGTGGATTGC GTCGGTCATG GTGCGCGGCC GTTTTCCCGA CACGGTTGAG  
 301 GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT  
 351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGTTGGAC AGACCCGGAA  
 401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT  
 451 TATGACGAAT TTTGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA  
 501 GATGACCTAT ACGGCACGTT CGGCCTGGAT TGTGTTTTG GACAACGGCA

264

```

551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTT
601 ACCGAAGCGT GGCAACATCT GTTGGGTAAA AATAAAATC GGTATCCTA
651 TGTGGATATG AGGTATAAGG ACGGATTTC AGTCCGCTAT GCTCCCGACG
701 GTTTACCCGA AAAAGAATCC GAAGAATAG

```

This corresponds to the amino acid sequence <SEQ ID 246; ORF 080.a>:

a080.pep

```

1 MWDNAEAMER LTRWLLVMM MLLAASGLVW FYNHNLVPK QVSLKGNLVY
51 SDKKALGSLA KEYIHGNILR TDINGAQEAY RYPWIASVM VRRRFPDTVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RGPMPVFRGA EGTSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY APDGLPEKES EE*

```

m080/a080 99.2% identity over a 242 aa overlap

	10	20	30	40	50	60
m080.pep	MWDNAEAMERL	TRWLLVMMAM	LLAASGLVWF	YNHNLVPKQV	SLKGNLVYSD	KKTLGSLA
a080	MWDNAEAMERL	TRWLLVMMAM	LLAASGLVWF	YNHNLVPKQV	SLKGNLVYSD	KKALGSLA
	10	20	30	40	50	60
	70	80	90	100	110	120
m080.pep	KEYIHGNILRT	DINGAQEAYR	YPWIASVMV	RRRFPDTVE	VVLTERKPVA	RWGDHALVDG
a080	KEYIHGNILRT	DINGAQEAYR	YPWIASVMV	RRRFPDTVE	VVLTERKPVA	RWGDHALVDG
	70	80	90	100	110	120
	130	140	150	160	170	180
m080.pep	EGNVFEARLDR	PGMPVFRGA	EGTSAEMLR	RYDEFSTVL	AKQGLGIKE	MTYTARSAW
a080	EGNVFEARLDR	PGMPVFRGA	EGTSAEMLR	RYDEFSTVL	AKQGLGIKE	MTYTARSAW
	130	140	150	160	170	180
	190	200	210	220	230	240
m080.pep	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLR	KNKNRLSY	VDMRYKDG	FSVRYASDGL
a080	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLR	KNKNRLSY	VDMRYKDG	FSVRYAPDGL
	190	200	210	220	230	240
m080.pep	EEX					
a080	EEX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 247>:

g081.seq

```

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51 GCCGTCTGAA AACAAACCCG TGTGCGCAT CGTAACCGAC AGCCGCGATA
101 TTCGGGAAGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGACCGG
151 CATGATTTTG TTGAGGCGT ATTGTCTGCG GCGCGGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGCGCGG CTTTGGGCGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC GTTGCAAACG TTGGCGAAGG CGTGGCGCGA TAATGTGAAC
301 CCGTTTGTGT TCGGCATTAC CGGTTGCGGC GGCAAGACGA CGGTGAAGGA
351 GATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTTTCGG
401 CGACGGCAGG CAACTTCAAC AACCAATcg gaTTGCCGCT GACTTTATTG
451 AAATtaacg aAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTTGGcgaa ctggcggtTt taacgcaaat CGCCAAACCC GATGCCGCTT
551 TGGtcaACAA CGCCTGCGC GCCCATGTCG GATGCGGTTt cgacggagtg
601 GGCGATATTG CCAAAGcgaa aagcGAGATT TatgcagGct tATGTTCAGA
651 CGGCATGSCA CTGATTCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGTT TAATTTGAAT ACGTGCACCT TCGGCGTCGA TAGCGGCGAT
751 GTCCGCGCGG AAAATATCGT GCTGAAACCT TTGTCGTGCG AATTGATTT

```



265

```

801 GGTGTGCGGC GACGAGCGCA CTGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GTCGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG CAAGGCTTCA GCAACATCAA
951 AGGCCGTCTG AACGTCAAAG CCGGCATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TCGGAATCCC GACAGTATGA AAGCCGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCCGCCGC CATTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAGGAc gaAGCCGCCG CCATGCACGC CGAagtcgGC GCGTACGCCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGcggaAA AATTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTTTGAGCC ACGATTGCC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTT GCGCTTTATG CAGAtggaAG AAGTgGTGGA GGCATTGGAG
1351 GATAAGTga

```

This corresponds to the amino acid sequence <SEQ ID 248; ORF 081.ng>:

g081.pep

```

1 MKPLDLNFC QALKLPMPSE NKPVSRIVTD SRDIREGDVF FALAGGRFDA
51 HDFVGGVLSA GAAAVVSRE DCAALGGALK VDDTLAALQT LAKAWRDNVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVSATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFE LAVLTQIAKP DAALVNNALR AHVGGCFDGV
201 GDIKAKSEI YAGLCSDGMA LIPQEDANMA VFKTATFNLN TCTFGVDSGD
251 VRAENIVLKP LSCEFDLVCG DERTAVVLPV PGRHNVHNA AALALALAAG
301 LSLNDVAEGL QGFSNIKRL NVKAGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMEGELGED EAAAMHAEV AYARDQGLEA AYFVGDSNVE
401 AAEKFGADGL WFAAKDPLIQ VLSHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 249>:

m081.seq

```

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51 GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGCGGAGCG GTTTGACCGC
151 CATGATTTTG TTGAAGACGT ATTGGCTGCT GGTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGTGCTG CAATGGATGG CCGGTTGAAA GTCGATGACA
251 CGCTTGCCGC ATTGCAAACG CTGGCAAAGG CGTGGCGTGA AAATGTGAAT
301 CCGTTTGTGT TCGGCATTAC CGGTTGCGGC GGCAAGACGA CGGTGAAGGA
351 AATGCTGGCT GCGGTATTGC GCCGCGGTTT CCGCGATGAT GCCGTGTTGG
401 CGACGGCAGG CAACTTCAAC AACCATATCG GATTGCCGCT GACTTTGTTG
451 AAGTTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTGCGCGAA CTGGCGGTTT TAACGCAMAT CGCAAACCA AATGCCGAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
601 GCGGATATTG CCAAAGCGAA AAGCGAGATT TACCAAGGTT TATGTTCAGA
651 CCGCATTGCA CTGATTCTC AAGAAGATGC CAATATGGCT GTCCTCAAAA
701 CCGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTTCACGCGG AAAATATTGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
801 GGTGTGCGGC GATGAGCGCG CCGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GTCGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGCG GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TCGGAACCCG GACAGCATGA AAGCTGCGAT TGACGTGTTG
1051 GCGCGTATGC CTGCCGCCGC TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAATG GCGGAGGACG AAGCCGCCG TATGCACGCC GAAGTCGGCG
1151 CGTATGCCCG CGACCAAGCG ATCGAAGCGG CTATTTTGTG CCGCGACAAC
1201 AGCGTCGAAG CGGCGGAAAA ATTTGGCGCG GACGGTTTGT GGTTCGCCGC
1251 CAAAGACCCG TTGATTCAAG TGTTGCGCCA CGATTTGCCG GAACGCGCCA
1301 CCGTGTGGT GAAAGGTTG CGCTTTATGC AGATGGAAGA AGTGGTCGAG
1351 GCATTGGAGG ATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 250; ORF 081>:

m081.pep

```

1 MKPLDLNFC QALKLPMPSE SKPVSVVTD SRDIRAGDVF FALAGERFDA
51 HDFVEDVLAA GAAAVVSRE DCAAMDGALK VDDTLAALQT LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFE LAVLTXIAKP NAALVNNAMR AHVGGCFDGV

```

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 081 shows 94.1% identity over a 455 aa overlap with a predicted ORF (ORF 081.ng) from *N. gonorrhoeae*:

m081/g081

	10	20	30	40	50	60
m081.pep	MKPLDLNFI	CQALKL	PMPSESK	PVSRIV	TD	SRDIRAGDV
g081	MKPLDLNFI	CQALKL	PMPSENK	PVSRIV	TD	SRDIREGDV
	10	20	30	40	50	60
m081.pep	GAAAVVVS	REDCAAM	DGALKV	DDTLAAL	QTLAKA	WRENVN
g081	GAAAVVVS	REDCAAL	GGALKV	DDTLAAL	QTLAKA	WRDENVN
	70	80	90	100	110	120
m081.pep	AVLRRR	FGDDAV	LATAGN	FNNHIG	LPLTLL	KLNEKH
g081	AVLRRR	FGDDAV	SATAGN	FNNHIG	LPLTLL	KLNEKH
	130	140	150	160	170	180
m081.pep	NAALVNN	AMRAHV	GC	FGDVG	VDIAKAK	SEIYQGL
g081	DAALVNN	ALRAHV	GC	FGDVG	VDIAKAK	SEIYAGLC
	190	200	210	220	230	240
m081.pep	TRTFGID	SGDVHA	ENIVLK	PLSCEF	DLVCGD	ERA
g081	TCTFGVD	SGDVRA	ENIVLK	PLSCEF	DLVCGD	ERT
	250	260	270	280	290	300
m081.pep	LSLNDVA	EGLKGF	SNIKGR	LNVKSG	IKGATL	IDDTYN
g081	LSLNDVA	EGLQGFS	NIKGR	LNVKAG	IKGATL	IDDTYN
	310	320	330	340	350	360
m081.pep	MGDMGEL	GELGEDE	AAAMHA	EVGAYAR	DQGIEA	AYFVGDN
g081	MGDMGEL	GE---	DEAAAM	HA	EVGAYAR	DQGIEA
	370	380	390	400	410	420
m081.pep	LIQVLR	HDLP	ERATV	LVKGS	RFMQME	EVVEA
g081	LIQVLS	HDLP	ERATV	LVKGS	RFMQME	EVVEA
	430	440	450			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 251>:

## a081.seq

```

1  ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51  GCCGTCTGAA AGCAAAACCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGATTGG CGGGCGGGCG GTTTGATGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCG GTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGCGTTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTACCGC GTTGCAAATG TTGGCGAAGG CGTGGCGCGA GAATGTGAAC
301 CCGTTTGTGT TCGGTATTAC CGGCTCGGGC GGCAAGACCA CGGTGAAGGA
351 AATGTTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATTTT GCCGTTTGG
401 CGACGGCAGG CAACTTCAAC AACCACATCG GATTGCCGTT GACTTTGTTG
451 AAATTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GTATGAACCA
501 TTTGGCGGAA CTGGCGGTTT TGACACAAAT CGCCAAACCC GATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
601 GGGGATATTG CCAAAGCGAA AAGCGAGAT TATCAAGCT TATGTTGAGA
651 CGGCATGGCG CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 GGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTCCACGCGG AAAATATCGT GCTGAAACCG TTGTCGTGCG AATTGATTG
801 GGTGTGCGGC AACGAGTGCG CAGCGTGGT TCTGCCCGTT CCCGCCCGCC
851 ACAATGTCCA CAACGCCGCC GCCGCCGCC CGCTGTCTTT GGCTGCAGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTT AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACCTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TCGGAACCTT GACAGCATGA AAGCTGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAAT
1101 GGGTGAGGAC GAAGCCGCCG CCATGCACGC CGAAGTCGGC GCGTACGCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGCAAA CAGCGTCGAA
1201 GCGGCGGAAA AATTTGCGCG GGACGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTTGCGCC ACGATTGGCC CGAACCGGCC ACCGTGTTGG
1301 TGAAAGGTTT CCGCTTTATG CAGATGGAAG AAGTGGTCCA GGCATTGGAG
1351 GATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 252; ORF 081.a>:

## a081.pep

```

1  MKPLDLNFIC QALKLMPSE SKPVSRIVTD SRDIRAGDVF FALAGGRFDA
51  HDFVEDVLAA GAAAVVVSRE DCVAMDGALK VDDTLTALQM LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRFRFGDN AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGELAVLTQIAKP DAALVNNAMR AHVGCDFDGV
201 GDIKAKSEI YQGLCSDGMA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251 VHAENIVLKP LSCEFDLVCG NECAAVVLPV PGRHNVHNAA AAAALS LAAG
301 LSNDVAEGL KGFSNIKGR LNVKSGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMLGELGED EAAAMHAEVG AYARDQGIEA AYFVGDN SVE
401 AAEKFGADGL WFAAKDPLIQ VLRHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

m081/a081 96.7% identity over a 455 aa overlap

	10	20	30	40	50	60
m081.pep	MKPLDLNFICQALKLMPSESKPVSRIVTDSDRDIRAGDVFFALAGERFDAHDFVEDVLAA					
a081	MKPLDLNFICQALKLMPSESKPVSRIVTDSDRDIRAGDVFFALAGGRFDAHDFVEDVLAA					
	10	20	30	40	50	60
m081.pep	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
a081	GAAAVVVSREDCAAMDGALKVDDTLTALQMLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
	70	80	90	100	110	120
m081.pep	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
a081	GAAAVVVSREDCAAMDGALKVDDTLTALQMLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
	70	80	90	100	110	120
m081.pep	AVLRFRFGDDAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTXIAKP					
a081	AVLRFRFGDNAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTQIAKP					

268

	130	140	150	160	170	180
m081.pep	190	200	210	220	230	240
a081	190	200	210	220	230	240
m081.pep	250	260	270	280	290	300
a081	250	260	270	280	290	300
m081.pep	310	320	330	340	350	360
a081	310	320	330	340	350	360
m081.pep	370	380	390	400	410	420
a081	370	380	390	400	410	420
m081.pep	430	440	450			
a081	420	430	440	450		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 253>:

```

g082.seq
1  atGTGGTTGT TGAAGTTGCC TGCCGTCGCC GAAACGGCAT CATCGCCGAA
51  ACGGCGGCGC AATACCGCAG CCAGCATCTC CTTACCGTC GTCTTGCCGC
101 CCGAACCGGT AATGCCGAAC ACAAACGGGT TCACATTATC GCGCCACGCC
151 TTCGCCAAGC TTTGCAACGC GGCAAGCGTG TCATCGACTT TCAACGCGCC
201 GCCCAAAGCC GCGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCGCCGC
251 CAGACAATAC GCCTCCAACA AAATCATGCG CGTCAAACCG CCCGCCCGCC
301 AATGCGAAAA ACACATCGCC TTCCCGAATA TCGCGGCTGT CGGTACGAT
351 GCGCGACACG GGTTTGTTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401 AGATGAAATT TAGGTCCAGT GGTTCATAT TTGCTTTCGT TAATATTCGG
451 GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501 GGTAACACAC AGATTATTTT CCCATTCTCA TTCGGCATT TTTCTGTACG
551 TATCATTTTT TAGACGTATT TTTAGCCGAT TTGCCTTTTC CCGCATACCA
601 CGGCGCGGGG TCGTCGGACT GTCTGTCGAT AAAGGCAAGG TTATTGCCTT
651 CGCCCGGCAC ATCGGGGACA TTCCCCCAA AATCATAGCC GTCATCGGGC
701 AACTCGTCGG TTCGATACC CGTCCAACG CCGAATCCGC GTAA

```

This corresponds to the amino acid sequence <SEQ ID 254; ORF 082.ng>:

```

g082.pep
1  MWLLKLPAVA ETASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTLSRHA
51  FANVCNAASV SSTFNAPPKA AQSSRETTTA AAPADNTPPT KSCASNRPPA
101 NAKNTSPSR1 SRLSVTMRDT GLFSDGIGSL RAWQMKFRSS GFIFAFVNIR
151 AADTSVAADF FIACFAVVKH RLFSSHSAF FLYVSFFRRI FSRFAFSRIP
201 RRGVGLSVD KGKVIAPARH IGDIPPKIIA VIGQLVGFD TPTAESAS*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 255>:

```

m082.seq
1  ATGnnGTTGT TGAAGTTGCC TGCCGTCGCC AACACGGCAT CATCGCCGAA
51  ACGGCGGCGC AATACCGCAG CCAGCATTTT CTTACCGTC GTCTTGCCGC
101 CCGAACCGGT AATGCCGAAC ACAAACGGAT TCACATTTTC ACGCCACGCC
151 TTTGCCAGCG TTTGCAATGC GGCAAGCGTG TCATCGACTT TCAACGCGCC

```

```
m082.pep
1  MXLLKLPAVA NTASSPKRRR NTAASISFTV VLPPEVPMNP TNGFTFSRHA
51  FASVCNAASV SSTFNAPSIA AQSSRETTTA AAPAANTSST KSCASNRPSP
101 NAKNTSPARM SRLSVTMRDT GLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151 AADTSVAADF FIACFAVAFH RLFSHSGSHX FLYVQFFRRR FSRFAFSRIP
201 RRGVVGOSVD TGKKVIAVLH IGNIPPKIIA VIGVLGVGFD TPTAESA*
```

ORF 082 shows 92.7% identity over a 247 aa overlap with a predicted ORF (ORF 082.ng) from *N. gonorrhoeae*:

		10	20	30	40	50	60
m082.pep		MXLLKLPAVANTASSPKRRRNTAASISFTVVLPEPVMPTNGFTFSRHAFASVCNAASV					
		:					
g082		MWLLKLPAVAETASSPKRRRNTAASISFTVVLPEPVMPTNGFTLSRHAFANVCNAASV					
		10	20	30	40	50	60
		70	80	90	100	110	120
m082.pep		SSTFNAPSIAAQSSRETTTAAAPAAANTSSTKSCASNRSpanakNTSPARMSRLSVTMRDT					
g082		SSTFNAPPKAAQSSRETTTAAAPADNTPPTKSCASNRPANAKNTSPSRISRLSVTMRDT					
		70	80	90	100	110	120
		130	140	150	160	170	180
m082.pep		GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF					
		:					
g082		GLFSDGIGSLRAWQMKFRSSGFIFAFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF					
		130	140	150	160	170	180
		190	200	210	220	230	240
m082.pep		FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT					
g082		FLYVSFFRRIFSRFAFSRIPRRGVVGLSVDKGKVIAFARHIGDIPKIIAVIGQLVGFDT					
		190	200	210	220	230	240
m082.pep		RPTAESAX					
g082		RPTAESAX					

1 ATGTGGTTGT TGAAGTTGCC TGCCGTCGCC AAAACGGCAT TATCGCCGAA  
51 ACGGCGGGCG AATACCGCAG CCAACATTTC CTTACCGTC GTCTTGCCGC

270

```

101 CCGAGCCGGT AATACCGAAC ACAAACGGGT TCACATTCTC GCGCCACGCC
151 TTCGCCAACA TTTGCAACGC GGTAAAGCGT TCATCGACTT TCAACGCGCC
201 ATCCATTGCA ACGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCCG
251 CAGCCAATAC GTCTTCAACA AAATCATGCG CATCAAACCG CCCGCCCGCC
301 AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTTACGAT
351 GCGCGACACG GGTTCGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGCG
401 AGATGAAATT TAJGTCCAGT GGTTCATAT TTTCTTCGT TAATATTCGG
451 GCGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCT GTTTTGCTGT
501 GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGCATT TTTCTGTACG
551 TATCATTTT TAGACGTATT TTTAGTCGAT TTGCCTTTT CCGCATACCA
601 CCGCGCGGGG TCGTCGGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
651 CGCCTGCAC ATCGGGAACA TTCCCCAAA AATCATAGCC GTCATCGGGC
701 AACTCGTCG TTTGATACC CGTCCAATG CCGAATCCG GTAA

```

This corresponds to the amino acid sequence <SEQ ID 258; ORF 082.a>:

a082.pep

```

1 MWLLKLPAVA KTALSPKRRR NTAANISFTV VLPPEPVI PN TNGFTFSRHA
51 FANICNAVSV SSTFNAPSIA TQSSRETTA AAPAANTSST KSCASNRP
101 NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151 AADTSVAADF FIACFAVVKH RLFSHSHSAF FLYVSFFRRI FSRAFSRIP
201 RRGVVQSV D KGVIAFALH IGNIPPKIIA VIGQLVGFD RPTAESAX

```

m082/a082 95.5% identity over a 247 aa overlap

	10	20	30	40	50	60
m082.pep	MXLLKLP	AVANTASS	PKRRRNTA	ASISFTV	VLPPEP	VMPNTNG
a082	MXLLKLP	AVAKTALS	PKRRRNTA	ANISFTV	VLPPEP	VI PN TNG
	10	20	30	40	50	60
m082.pep	SSTFNAPS	IAAQSSRE	TTTAAAPA	ANTSSTK	SCASNRS	PANAKNTS
a082	SSTFNAPS	IAAQSSRE	TTTAAAPA	ANTSSTK	SCASNRP	PANAKNTS
	70	80	90	100	110	120
m082.pep	PARMSRL	SVTMRDT				
a082	PARMSRL	SVTMRDT				
	70	80	90	100	110	120
m082.pep	GLLSDGIG	SLRAWQMK	FRSSGFIF	TFVNIRA	ADTSVAAD	FFIACFAV
a082	GLLSDGIG	SLRAWQMK	FRSSGFIF	TFVNIRA	ADTSVAAD	FFIACFAV
	130	140	150	160	170	180
m082.pep	VRHLFSH	SHSXF				
a082	VRHLFSH	SHSXF				
	130	140	150	160	170	180
m082.pep	FLYVSFFR	RIFSRFA	FSRIPRR	GVVQSV	DKGVIA	FALHIGNI
a082	FLYVSFFR	RIFSRFA	FSRIPRR	GVVQSV	DKGVIA	FALHIGNI
	190	200	210	220	230	240
m082.pep	PPKIIA	VIGQLV	GFD			
a082	PPKIIA	VIGQLV	GFD			
	190	200	210	220	230	240
m082.pep	RPTAESAX					
a082	RPTAESAX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 259>:

g084.seq

```

1 ATGAAacaAT CCGcccgaaT aAAAAATATG GATCAGACAT TAAAAAATac
51 attgggcatt tGCGCGcttt tagcctTTTG TTTggcgG gccaTCGCAT
101 CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGC
151 GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GCTTCCCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTACGT CGGCACAACC GCCCTATATT
251 TGCCGGTCGG CTGGCTGTAT GGTGCGCCTT CTTATCAGAT AGTCGGTTCC
301 ATATTGGAAG GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
351 CGGGTGCGCTT TATTTTGTGC AGGCATTATT TTTCATTTTT GGCTTGACAG

```

271

```

401 TTTGAAATA TTGTGTATCT GTGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGTCGCA GCAAAATATG GCTGACCATA TTATTGACTT TGATTTTGTC
501 CTGCGCGGTG ATGGAGAAAA TCGccggcga taaAGATTGG CGAGaacctg
551 atgccggcct gttgttgaat ATTTTcgacc tgtattaCga cttggctttc
601 cgcgccggca cAATATGCCG CCAAGCGCGC CCacattttg gaagCagcaa
651 aaaaagcgtC AACATGGCAt atccgccaac ttgcgcccaa gTAtaa

```

This corresponds to the amino acid sequence <SEQ ID 260; ORF 084.ng>:

```

g084.pep
1  MKQSARIKMN DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
51  ALASVVFLLL LARGFPRVSS VLLIYVGT ALYLPVGWLY GAPSYQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS VGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MEKIAGDKDW REPDAGLLN IFDLYYDLAF
201 RAGTICRQAR PHFGSSKKSVMAYPPTCAQ V*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 261>:

```

m084.seq
1  ATGAAACAAT CCGCCcGAAr AAAa.ATATG AATCAGACAT TACTTTATAC
51  ATTGGGCATT TGCGCGCTTT TAACCTTTnn nnnnnnnnnn nnnnnnnnnn
101 nnnnnTATCA CCCnGAATAT GAATACGGCT ACCGTTATTC TGCGTGGGT
151 GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GTTCCCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTACGT CGGCACAACC GCCCTATATT
251 TGCCGTCGG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
301 ATATTGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
351 CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCATTTT GGCTTGACAG
401 TTGGAATA TTGTGTATCG GGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTC
501 CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTG CGAGAACCTG
551 ATGCCGGCCT GTTGTGAAT ATTTTCGACC TGTATTACGA TTTGGCT.TC
601 CGCGCCGCA CAATATGCCG CCAAGCGCGC CCACATTTG GAAGCAGCAA
651 AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA

```

This corresponds to the amino acid sequence <SEQ ID 262; ORF 084>:

```

m084.pep
1  MKQSARIKXM NQTLTYTLGI CALLTFXXXX XXXXXHPEY EYGYRYSAVG
51  ALASVVFLLL LARGFPRVSS VLLIYVGT ALYLPVGWLY GAPSYQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS GGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLN IFDLYYDLAX
201 RAGTICRQAR PHFGSSKKSVMAYPSCCAQ V*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 084 shows 90.5% identity over a 231 aa overlap with a predicted ORF (ORF 084.ng) from *N. gonorrhoeae*:

```

m084/g084
10      20      30      40      50
m084.pep MKQSARIKXMNQTLTYTLGICALLTF-----YHPEY EYGYRYSAVGALASVVFLLL
||||||| :||| |||||:| || ||||| ||||| |||||
g084      MKQSARIKXMDQTLKNTLGICALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFLLL
10      20      30      40      50      60
60      70      80      90      100     110
m084.pep LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
||||||| :||| |||||:| ||||| ||||| ||||| |||||
g084      LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
70      80      90      100     110     120
120     130     140     150     160     170
m084.pep YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL
||||||| :||| |||||:| ||||| ||||| ||||| |||||
g084      YFVQALFFIFGLTVWKYCVSVGVFADVKNYKRRSKIWLTI LLTLILSCAVMEKIAGDKDW
130     140     150     160     170     180

```

272

	180	190	200	210	220
m084 .pep	REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVMAYPSCCAQVX				
g084	REPDAGLLLNIFDLYYDLAFRAGTICRQARPHFGSSKKSVMAYPPTCAQVX				
	190	200	210	220	230

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 263>:

**a084 .seq**

```

1  ATGAAACAAT CCGCCCGAAT AAAAAATATG GATCAGACAT TAAAAAATAC
51  ATTTGGGCATT TGGCGCTTT TAGCCTTTTG TTTTGGCGCG GCCATCGCAT
101 CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTATTTC TGCCGTGGGT
151 GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GTTCCCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251 TGCCGCTCGG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
301 ATATTGAAA GCAATCCTGC CGAGGCGCGT GAATTGTGCG GCAATCTTCC
351 CGGGTCGCTT TATTTGTGC AGGCATTATT TTTCATTTT GGCTTGACAG
401 TTTGGAGATA TTGTGTATCG GGGGGGTAT TTGCTGACGT AAAAACTAT
451 AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTG
501 CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTG CGAGAACCTG
551 ATGCCGCGCT GTTGTGAAT ATTTTCGACC TGTATTACGA TTTGGCTTCC
601 .GCGCCGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
651 AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA

```

This corresponds to the amino acid sequence <SEQ ID 264; ORF 084.a>:

**a084 .pep**

```

1  MKQSARIKNM DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
51  ALASVVFLLL LARGFPRVSS VLLIYVGT ALYLPVGWLY GAPS YQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWRYCVS GGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLN IFDLYYDLAS
201 XAGTICRQAR PHFGSSKKSVMAYPSCCAQ V*

```

**m084/a084** 92.2% identity over a 231 aa overlap

	10	20	30	40	50	60
m084 .pep	MKQSARIKXMNQTLTYLTGICALLTFXXXXXXXXXXHPEYGYRYSAVGALASVVFLLL					
	:          :					
a084	MKQSARIKNMDQTLKNTLGIALLAFCFGAIAISGYHLEYEYGYRYSAVGALASVVFLLL					
	10	20	30	40	50	60
m084 .pep	LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILES NPAEAREFVGNLPGSL					
a084	LARGFPRVSSVLLIYVGTALYLPVGWLYGAPSYQIVGSILES NPAEAREFVGNLPGSL					
	70	80	90	100	110	120
m084 .pep	YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL					
a084	YFVQALFFIFGLTVWRYCVSGGVFADVKNYKRRSKIWLTI LLTLILSCAVMDKIASDKDL					
	130	140	150	160	170	180
m084 .pep	REPDAGLLLNIFDLYYDLAXRAGTICRQARPHFGSSKKSVMAYPSCCAQVX					
a084	REPDAGLLLNIFDLYYDLASXAGTICRQARPHFGSSKKSVMAYPSCCAQVX					
	190	200	210	220	230	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 265>:

**g085 .seq**

```

1  ATGGGCAAAG GGCAGGACTT CAGCCCCCTG CGCGACGCGT TGAAAGATAA

```



273

```

51  GGCAAAAGGC GTGTTCTGA TCGGCGTCGA TCGGCCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AACCTGACCG ACTGCGTCAC TTGGAAGAG
151 GCGGTTTCAGA CGGCATACGC CCAAGCCGAA GCGGGCGATA TTGTCTTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCCTT
251 CGGAAGTGTt tatCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 266; ORF 085.ng>:

```

g085.pep
1  MGKGQDFTPL RDALKDKAKG VFLIGVDAPO IRRDLGCGL NLTDCVTLEE
51 AVQTAYAQAe AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 267>:

```

m085.seq
1  ATGGGTAAAG GGCAGGACTT CACGCCCTG CGCGATGCAC TGGTAGGCAA
51 GGCAAAAGGC GTGTTCTGA TTGGTGTGCA TCGGCCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AATATGACCG ACTGCGCCAC TTGGGAGAA
151 GCGGTTTCAGA CGGCATATGC CCAAGCCGAA GCAGGCGATA TTGTGTGCT
201 CAGCCCCGCC TCGCGAGCT TTGATATGTT CAAAGGCTAC GCGCACCCTT
251 CGGAAGTGTt TATCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 268; ORF 085>:

```

m085.pep
1  MGKGQDFTPL RDALVGKAKG VFLIGVDAPO IRRDLGCGL NMTDCATLGE
51 AVQTAYAQAe AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 085 shows 94.7% identity over a 94 aa overlap with a predicted ORF (ORF 085.ng)

from *N. gonorrhoeae*:

m085/g085

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGLNMTDCATLGEAVQTAYAQAe					
g085	MGKGQDFTPLRDALKDKAKGVFLIGVDAPOIRRDLGCGLNLTDCVTLEEAVQTAYAQAe					
	10	20	30	40	50	60
	70	80	90			
m085.pep	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
g085	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
	70	80	90			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 269>:

```

a085.seq
1  ATGGGCAAAG GGCAGGACTT CACGCCCTG CGCGACGCGC TTGCCGGCAA
51 GGCAAAAGGC GTGTTCTGA TCGGTGTCGA TCGGCCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGATCTG AATATGACCG ACTGCGCCAC TTGGAAGAA
151 GCGGTTTCAGA AGGCATATGC CCAAGCCGAA GCGGGCGATA TCGTGCTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCCTT
251 CGGAAGTGTt TATCGGGGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 270; ORF 085.a>:

```

a085.pep
1  MGKGQDFTPL RDALAGKAKG VFLIGVDAPO IRRDLGCDL NMTDCATLEE
51 AVQKAYAQAe AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

m085/a085 94.7% identity over a 94 aa overlap.

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGLNMTDCATLGEAVQTAYAQAe					

```

a085      |||||:||||'|||||| ||||| ||||| |||||
           10      20      30      40      50      60
m085.pep  AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
           70      80      90
a085      AGDIVLLSPACASFDMFKGYAHRSEVFIGAFKALX
           70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 271>:

```

g086.seq
1  ATGGTGGTGC TGATGACGGC GTTCGGCCTG CTGATGATTT ATTCGGCTTC
51  TGTGTATTTG GCATCGAAGG AAGGCGGCGA TCAGTTTTC TATTTGACCA
101 GGCAGGCGGG GTTCGTCGTT GCCGGCCTTA TAGCGAGCGG TTTTATATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTCG
201 CTTATCCGGC CTGTTGCTGG TAGCCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTTC AAGCTGGCAGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
351 CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
401 GGCGGGGGAC GGCCAACCTG ATTATGTCCG CCACCAATCC GCAGGCACGT
451 CGTGAACAT TAGAAATGTA CGGcCGTTTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTTCG
551 GTTCGTTTGT CGTCATTACC GTCATTACCG TTGGAATGCT GTTCTGCGCA
601 GGATTGCCGT GGAAATATTT TTTGTCTCTG GTAGGCGAGC TCTTGGGTGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTGGGA CCCGTGGAAA GACCCGAGG GTGCCGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGTA TGGGTTTGGG
801 TGCGAGTTTG AGCAAACGCG GCTTCTGCCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCTGAA GAATTCGGCT TCTTCGGGAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGtttgA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGGAT CCGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA CCAAAGGTCT GACGctgCcg tTGATGTCCT ATGGcggTTC
1101 GTCAGTCTTT TCATGCTGTA TCAGCATGAT GCTGCTGTTG CGTATCGATT
1151 ATGAAAACCG CCAGAAAATG CGCGGTTACC GGGTGGAGTA AA

```

This corresponds to the amino acid sequence <SEQ ID 272; ORF 086.ng>:

```

g086.pep
1  MVVLMTAFGL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGFLW
51  FLCRMRTWRR LVPWIFALSG LLLVAVLIAG REINGATRWI PLGPLNFQPT
101 ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQAR
151 RETLEMYGRF RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VITVGMLFLA
201 GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAPLDPWK LPQGAGYQLT
251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301 IFCYGWLVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGIQ SFFNIGVNIG
351 ALPTKGLTLP LMSYGGSSVF FMLISMMLLL RIDYENRQKM RGYRVE*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 273>:

```

m086.seq
1  ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
51  TGTGTATTTG GCATCAAAAAG AAGGCGGCGA TCAGTTTTC TATTTGACCA
101 GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTCG
201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTTC AAGCTGGCGGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
351 CCGTGAAGAA GTGTTGcGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
401 GGCGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGrCACGT
451 CGTGAaACAT TAGAAATGTA CGGCCGTWTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTTCG

```

```

551 GTTCGTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTGGCA
601 GGATTGCCGT GGAAATATTT TTTCGTCCTG GTAGGCAGCG TCTTGGGCGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTTGGA CCCGTGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801 TGGGAGTTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
851 TTTTGGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGTTTGA CTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGGAT CCGkTCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA mCAAAGyCT GACGCyGCCG Tg.AGTCCW ATGGCGGTTT
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTkG CGTATAGATT
1151 ATGAAAACCG CCGGAAAATG CGCGTTATC GGGTGGAGTA A

```

This corresponds to the amino acid sequence <SEQ ID 274; ORF 086>:

```

m086.pep
1  MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
51  FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINCATRWI PLGPLNPQPT
101 ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQXR
151 RETLEMYGRX RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMLFLA
201 GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFDPWK DPQAGYQLT
251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301 IFCYGWLVRV AFSIGKQSRD LGITFNAYIA SGIGIWIGXQ SFFNIGVNI
351 ALPXKGLTXP XMSXGSSVF FMLISMMLLX RIDYENRRKM RGYRVE*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 086 shows 96.7% identity over a 396 aa overlap with a predicted ORF (ORF 086.ng)

from *N. gonorrhoeae*:

```

m086/g086
      10      20      30      40      50      60
m086.pep  MVVLMTAFSLLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
g086      MVVLMTAFGLLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLVWFLCRMRTWRR
      10      20      30      40      50      60

      70      80      90     100     110     120
m086.pep  LVPWIFALSGLLLVVLIAGREINGATRWIPLGPLNPQPTLFLKAVILYLASLFTRREE
g086      LVPWIFALSGLLLVAVLIAGREINGATRWIPLGPLNPQPTLFLKAVILYLASLFTRREE
      70      80      90     100     110     120

      130     140     150     160     170     180
m086.pep  VLRSMESLGWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAIILPIMLVAFGLVLIMVQ
g086      VLRSMESLGWQSIWRGTANLIMSATNPQARRETLEMYGRFRAIILPIMLVAFGLVLIMVQ
      130     140     150     160     170     180

      190     200     210     220     230     240
m086.pep  PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQVVAFDPWK
g086      PDFGSFVVITVITVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQVVAFDPWK
      190     200     210     220     230     240

      250     260     270     280     290     300
m086.pep  DPQAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL
g086      DPQAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIEEFGFFGMCVL
      250     260     270     280     290     300

      310     320     330     340     350     360

```

276

```
m086.pep      IFCYGWLVRAPFSIGKQSRDLGLTFNAYIASGGIGIWIGXQSFFNIGVNIGALPXKGLTXP
|||||
g086          IFCYGWLVRAPFSIGKQSRDLGLTFNAYIASGGIGIWIGIQSFFNIGVNIGALPTKGLTLP
              310       320       330       340       350       360

              370       380       390
m086.pep      XMSXGGSSVFFMLISMMLLXRIDYENRRKMRCGYRVEX
|||
g086          LMSYGGSSVFFMLISMMLLRIDYENRQKMRGYRVEX
              370       380       390
```

a086.seq

1	ATGGTGGTG	CGATGACGGC	GTTGAGCGTG	CTGATGATT	ATTCGGGCTC
51	TGTGATTGG	GCATCAAAAG	AAGGCGGCGA	TCAGTTTTC	TATTTGACCA
101	GACAGGCGGG	GTTGCTCGTT	GCGGGCTTGA	TAGCAGAGCG	TGTGTTATGG
151	TTTCTTTGCA	GGATGAGGAC	ATGGCGGCGG	CTGTGCCGT	GGATTTTTCG
201	CCTATCCGGC	CTGTTGCTGG	TAGTCGTATT	GATTGCCGGG	CGCGAAATCA
251	ATGGCGCGAC	CCGTTGGATA	CCTTTGGGTC	CGTTGAATT	CCAGCCGACG
301	GAGCTGTTCA	AGCTGGCGGT	CATCTTTTAT	TTGGCAAGCC	TGTTCTACCG
351	CCGTGAAGAA	TGCTTGCGCA	GATGGAAGG	TTGGGTTGG	CAGTCGATT
401	GGCGGGGGAC	GGCCAATCTG	ATCATGTCCG	CCACCAATCC	GCAGGCACGT
451	CGTGAACAT	TAGAAATGTA	CGGCGGTTT	CGGGCGCAT	TCCTGCCGAT
501	TATGCTGGTG	CGCTTCGGTT	TGGTCTGTAT	ATAGGTACAG	CCGGATTTCG
551	GTTCTTTGT	CGCTATTACC	GTTCGTCGG	TTGGAATGCT	GTTTTTGCCA
601	GGATTGCCGT	GGAATAATTT	TTTCGTCCTG	GTAGGCAGCG	TCTTGGGCGG
651	GATGGTGCTG	ATGATTACCG	CGGCTCCCTA	CCGTTGTGAC	CGGGTAGTGG
701	CATTTTGGGA	CCCGTGGAAA	GACCCGACAG	GTCGCCGCTA	CCAGCTTACC
751	CACCTCTGTA	TGGCAATCGG	GCGCGGAGAG	TGCTTCGTA	TGGGTTTGGG
801	TGCGAGTTTG	AGCAAACGCG	GCTTCTGCTC	GGAAGCGCAT	ACCGATTTTA
851	TTTTTGCCAT	CATPCGCCAA	GAATTCGGTT	TCTTCGGTAT	TGCGGTGCTG
901	ATATCTGTT	ACGGCTGGCT	GGTGGTGGCG	CGCTTTTCCA	TCGGCAAGCA
951	GTCGCGCGAT	TTGGGTTTGA	CTTCAACGCG	CTATATCGCT	TCGGGTATCG
1001	GCAATTGGAT	CGGTATCCAA	AGTTTCTTCA	ATATCGGTGT	GAACATCGGT
1051	GCTTTGCCGA	CCAAAGGCTC	GACGCTGCCG	TGTATGTCCT	ATGGCGGTTT
1101	GTCAGTCTTT	TTCACTGCTA	TCAGCATGAT	CTGCTGGTTG	CGTATAGATT
1151	ATGAAAACCG	CCGGAATAATG	CGCGGTTACC	GGGTGGAGTA	A

a086.pcp

1	MVVLMTAFSL	LMIYSASVYL	ASKEGGDQFF	YLTRQAGFVV	AGLIASGLLW
51	FLCRMRTWRR	LVPWIFALSG	LLLVVSLIAG	REINGATRWI	PLGPNLFQPF
101	ELFKLAVILY	LASLFTTRSE	VLRSMVSLGW	QSIWRGTANL	IMSATNPOAR
151	RETELMYGRF	RAILPLMLV	AFGLVLMIVQ	PDFGSFVVIT	VIAGVLMFLA
201	GLPWKYFFVL	VGSVLGGMVL	MITAAPRYVQ	RVVAFLDPWK	DPQGAGYQLT
251	HSLMAIGRGE	WFGMGKQSL	SKRGFLPEAH	TDFIFAIIE	EFGFGMGCVL
301	IFCYGWLVR	AFSIGKQSRD	LGLTFNAYIA	SGIGIWIQAE	FFNIGMNVIG
351	ALPTKGLTLP	LMSYGGSSVF	FMLISMMLLL	RIDYENRRRK	RGYRVE*

m086/a086 98.0% identity over a 396 aa overlap

```

      10      20      30      40      50      60
m086.pep  MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
          |||||
a086      MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR
          |||||
      10      20      30      40      50      60

      70      80      90     100     110     120
m086.pep  LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTTEFLKAVILYLASLFTTRREE
          |||||
a086      LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTTEFLKAVILYLASLFTTRREE
          |||||

```

277

	70	80	90	100	110	120
	130	140	150	160	170	180
m086.pep	VLRSMESLGWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAIILPIMLVAFGLVLMVQ					
a086	VLRSMESLGWQSIWRGTANLIMSATNPQARRETLEMYGRFRAIILPIMLVAFGLVLMVQ					
	130	140	150	160	170	180
	190	200	210	220	230	240
m086.pep	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
a086	PDFGSFVVITVIAVGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
	190	200	210	220	230	240
	250	260	270	280	290	300
m086.pep	DPQGAGYQLTHSLMAIGRGWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFGFFGMCVL					
a086	DPQGAGYQLTHSLMAIGRGWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFGFFGMCVL					
	250	260	270	280	290	300
	310	320	330	340	350	360
m086.pep	IFCYGWLVRFAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNI GALPXGLTYP					
a086	IFCYGWLVRFAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNI GALPTGLTLP					
	310	320	330	340	350	360
	370	380	390			
m086.pep	XMSXGGSSVFFMLISMMLLXRIDYENRRKMRGYRVEX					
a086	LMSYGGSSVFFMLISMMLLLRIDYENRRKMRGYRVEX					
	370	380	390			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 277>:

g087.seq

```

1  ATGGGCGGTA  AAACCTTTAT  GCTGATGGCG  GCGGAACGG  GCGGACACAT
51  TTTCCAGCT   CTGGCTGTGG  CGGATTCATT  GCGCGTGC GC GTCATCATG
101 TAATTTGGCT  GGGCAGCAAG  GATTCGATGG  AAGAGCGCAT  CGTGCCGCAA
151 TACGGCATA C GCTTGGAAAC  GCTGGCGATT  AAAGGAATAC  GCGGCAACGG
201 CATCAAACGC  AAGCTGATGC  TTCCGTTTAC  TCTGTACAAA  ACCGTCGCGC
251 AAGCGCAGCG  GATTATCCGC  AAACACCGTG  TCGAGTGCCT  CATCGGCTTC
301 GCGGTTTTG  TTACCTTTCC  CGGCGGTCTG  GCGGCGAAAC  TCTTGGGCGT
351 GCCGATTGTG  ATTCACGAGC  AAAACGCCGT  GGCAGGCTTG  TCCAACCGCC
401 AccTGTCGCG  ctGGGCGAAA  CGGGTGTTGT  ACCTTTTCC  GAAAGCGTTC
451 AGCCACGAAG  GCGGTTTTGT  CGGCAACCCC  GTCCGCGCCG  ATATTAGCAA
501 CTTGCCCCGT  CCTGCCGAAC  GCTTCCAAGG  GCGCGAAGGC  CGTCTGAAAA
551 TTTTGGTGGT  CGGCGGCAGT  TTGGGTGCGG  ACGTTTTGAA  CAAAACCGTA
601 CCGCAGGCGT  TGGCACTGCT  GCCTGAAGAG  GTGCGCCCGC  AGATGTACCA
651 CAGTCGGGG  CGTAACAAGC  TGGGCAATCT  TCAGGCGGAT  TATGACGCGT
701 TGGGCGTGAA  AGCGGAATGC  GTGGAATTTA  TTACCGACAT  GGTGTCCGCC
751 TACCGTGATG  CCGATTTGGT  GATTTGCCGT  GCCGGCGCGC  TGACGATTGC
801 CGAGTTGACG  GCGGCGGGGC  TGGGCGCGTT  GTTAGTGCCG  TATCCTCACG
851 CCGTTGATGA  CCATCAAACC  GCCAACGCGC  GTTTCATGGT  GCAGGCAGAA
901 GCGGGGCTGC  TGTTGCCGCA  AACCAGTTG  ACGGCGGAAA  AACTCGCCGA
951 AATCCTCGGC  AGCCTCAACC  GCGAAAAATG  CCTCAAATGG  GCGGAAAACG
1001 CCCGTACGTT  GGCAATGCCG  CACAGCGCGG  ATGACGTGCG  CGAAGCCGCG
1051 ATTGCGTGTG  CGGCGTAAA

```

This corresponds to the amino acid sequence <SEQ ID 278; ORF 087.ng>:

g087.pep

```

1  MGGKTFMLMA  GGTGGHIFPA  LAV  LSLRVR  GHVIVLWLSK  DSMEERIVPQ
51  YGIRLET LAI  KGIRNGIKR  KLMLPFTLYK  TVREAQRIIR  KHRVECVIGF
101 GGFVTFPGGL  AAKLLGVPIV  IHEQNAVAGL  SNRHLSRWAK  RVLVAFPKAF
151 SHEGGLVGNP  VRADISNLPV  PAERFQGGREG  RLKILVVGGS  LGADVLNKTV

```

201 PQALALLPEE VRPQMYHOSG RNKLGNLQAD YDALGVKAEC VEFITDMVSA  
 251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE  
 301 AGLLLPQTQL TAEKLAEILG SLNREKCLKW AENARTLALP HSADDVAEAA  
 351 IACAA\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 279>:

m087.seq  
 1 ATGGGCGGTA AAACCTTTAT GCTGAWkkCG GCGGGAACGG GCGGACATAT  
 51 TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG  
 101 TGATTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGTAT CGTGCCGCAA  
 151 TACGGCATACT GCTTGGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG  
 201 CATCAAACGC AAAGTATGTC TGCCGGTTAC TTTGTATCAA ACCGTCCGCG  
 251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCCT CATCGCCTTC  
 301 GCGGCGTTTCG TTACCTTCCC CGGCGGTTTG GCGGCGAAGC TATTArGCGT  
 351 GCGGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC  
 401 ACCTGTCCGC CTGGGCGAAG CGGGTGTGTG ACGCTTTTCC GAAAGCGTTC  
 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA  
 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA  
 551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA  
 601 CCGCATGCAT TGGCTTTGCT GCCGACAAT GCGCGTCCGC ATATGTACCA  
 651 CCAATCGGGA CGGGGCAAGC TGGGCATCTT GCAGGCGnnn nnnnnnnnnn  
 701 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn  
 751 nnnGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCAGC CGGTTGACGA  
 801 TCACCAAACC GCCAACGCGC GTTTATGGT GCAGGCGGAG GCGGGATTGC  
 851 TGTGCGCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA GATTCTCGGC  
 901 GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG CCCGTACGTT  
 951 GGCAGTCCG CACAGTGCGG ACGACGTGGC GGAAGCCGCG ATTGCGTGTG  
 1001 CGGCGTAA

This corresponds to the amino acid sequence <SEQ ID 280; ORF 087>:

m087.pep  
 1 MGGKTFMLXX GGTGGHIFPA LAVADSLRAR GHVIVLGSK DSMEERIVPQ  
 51 YGIRLETIAI KGVIRNGIKR KLMLPVTLYQ TVREAQRIIR KHRVECVIGF  
 101 GGFVTFPGGL AAKLLXVPIV IHEQNAVAGL SNRHLSRWAK RVLVAFPKAF  
 151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGS LGADVNLKTV  
 201 PHALALLPDN ARPHMYHOSG RGKLGILQAX XXXXXXXXXX XXXXXXXXXX  
 251 XAGLGALLVP YPHAVDDHQT ANARFMVQAE AGLLLPQTQL TAEKLAEILG  
 301 GLNREKCLKW AENARTLALP HSADDVAEAA IACAA\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 087 shows 83.9% identity over a 355 aa overlap with a predicted ORF (ORF 087.ng) from *N. gonorrhoeae*:

m087/g087

	10	20	30	40	50	60
m087.pep	MGGKTFMLXXGGTGGHIFPALAVADSLRARGHVVIVLGSKDSMEERIVPQY	YGIRLETIAI				
g087	MGGKTFMLMAGGTGGHIFPALAVADSLRVRGHVVIVLGSKDSMEERIVPQY	YGIRLETIAI				
	10	20	30	40	50	60
	70	80	90	100	110	120
m087.pep	KGVIRNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPIV					
g087	KGIRNGIKRKLMLPFTLYKTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m087.pep	IHEQNAVAGLSNRHLSRWAKRVLVAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
g087	IHEQNAVAGLSNRHLSRWAKRVLVAFPKAFSHEGGLVGNPVRADISNLPVPAERFQREG					
	130	140	150	160	170	180

279

```

          190      200      210      220      229
m087.pep  RLKILVVGSSLGADVLNKTVPHALALLPDNARPHMYHQSGRGKLGILQA-----
          |||||||
g087      RLKILVVGSSLGADVLNKTVPQALALLPEEVPRQMYHQSGRNKLGNLQADYDALGVKAEC
          190      200      210      220      230      240

          230      240      250
m087.pep  -----AGLGALLVPYPHAVDDHQTANARFMVQAE
          |||||||
g087      VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE
          250      260      270      280      290      300

          260      270      280      290      300      310
m087.pep  AGLLLPQTQLTAEKLAELGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
          |||||||
g087      AGLLLPQTQLTAEKLAELGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
          310      320      330      340      350

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 281>:

```

a087.seq
1   ATGGGCGGTA AAACCTTTAT GCTGATGGCG GCGGGAACGG GCGGACATAT
51  TTTCCCGCGC CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
101 TAATTGGGCT GGGCAGCAAG GATTTCGATG AAGAGCGCAT CGTGCCGCAA
151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG
251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
301 GCGCGCTTCG TTACCTTTCC GCGCGGTTTG GCGGCGAAGT TATTAGGCGT
351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC
401 ACCTGTCGCG CTGGCGGAAG CGGGTGTGTG ACGCTTTTCC GAAAGCGTTC
451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCGG ATATTAGCAA
501 CCTGCCCGTG CTGCGCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAFA
551 TTTTGGTGGT CGGCGGCACT TTGGGCGCGG ACGTTTGTAA CAAAACCGTA
601 CCGCAGGCAT TGGCTTTGCT GCGCGACAAT GCGCGTCCGC AGATGTACCA
651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC
701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC
751 TACCGCGATG CCGATTGGT GATTTGCCGT GCCGCGCGCG TGACGATTGC
801 CGAGTTGACG GCGGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCAGC
851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGTT GCAGGCGGAG
901 GCGGGATTGC TGTGCGCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA
951 GATTCTCGGC GGCTTAAACC GCGAAAATG CCTCAAATGG GCAGAAAACG
1001 CCCGTACGTT GGCAGTCCG CACAGTCCG ACGAGCTTGC CGAAGCGCGC
1051 ATTGCGTGTG CGGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 282; ORF 087.a>:

```

a087.pep
1   MGGKTFMLMA GGTGGHIFPA LAVADSLRAR GHHVIWLGSK DSMEERIVPQ
51  YDILLETLAI KGVVRNGIKR KLMLPFTLYQ TVREAQQIIR KHRVECVIGF
101 GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLVAFPKAF
151 SHEGGLVGNP VRADISNLPV PAERFQGREG RLKILVVGGS LGADVLNKTVP
201 PQALALLPDN ARPQMYHQSG RGKLGSLQAD YDALGVQAEC VEFITDMVSA
251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHOT ANARFMVQAE
301 AGLLLPQTQL TAEKLAELG GLNREKCLKW AENARTLALP HSADDVAEAA
351 IACAA*

```

m087/a087 85.4% identity over a 355 aa overlap

```

          10      20      30      40      50      60
m087.pep  MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYIGIRLETLAI
          |||||||
a087      MGGKTFMLMAGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYDILLETLAI
          10      20      30      40      50      60

```

280

m087.pep	70	80	90	100	110	120
	KGVRGNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVP					
a087	70	80	90	100	110	120
	KGVRGNGIKRKLMLPFTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV					
m087.pep	130	140	150	160	170	180
	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
a087	130	140	150	160	170	180
	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
m087.pep	190	200	210	220	230	240
	RLKILVVGGSLGADVLNKTVPHALALLPDNARPHMYHQSGRGKLGILQAXXXXXXXXXXX					
a087	190	200	210	220	230	240
	RLKILVVGGSLGADVLNKTVPQALALLPDNARPMYHQSGRGKLGSLQADYDALGVQAEC					
m087.pep	250	260	270	280		
	XX-----XXXXXXXXXAGLGALLVPYPHAVDDHQTANARFMVQAE					
a087	250	260	270	280	290	300
	VEFITDMVSAYRDADLVICRAGALTAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE					
m087.pep	290	300	310	320	330	
	AGLLLPQTQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
a087	290	300	310	320	330	
	AGLLLPQTQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 283>:

g088.seq

```

1  ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
51  TTTTCAATAC ACCACATTCC GCGCCGTTAT GGCGGCGTTG ACCGCCTTGG
101 CGTTTTCCCT GATGTTTCGGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
151 AAATGCGGGC AGGCAGTGCG TACCGACGGC CCGCAAACCC ACCTCGTCAA
201 AAACGGCACG CCGACGATGG GCGGTTGCTT GATTCTGACC GCCATTACCG
251 TGTCCACCCCT GTTGTGGGGC AACTGGGCGA ACCCGTATAT CTGGATTCTC
301 TTGGGCGTAC TGCTTGCCAC CGGTGCGCTC GGTTTTTACG ACGACTGGCG
351 CAAAGTCGTT TATAAAGACC CCAACGGCGT GTCGCGCAAA TTCAAATGG
401 TGTGGCAGTC AAGCGTTGCC GTTatcgCG GTttggcatt GTTTTACctt
451 gCcgCcaATT CCGCCAACAA TATTTTGATT GTCCCGtttT TCAAACAAAT
501 CGCCCTGCCG CTGGGCGTGG TCGGCTTttt gGtggtgTCT TACCTGACCA
551 TCGTCGGCAC ATCCAACGCC GTCAACCTCA CgaCGGCTT GGACGGCCTT
601 GCGGCcttcc cgttcgtcct cgttgccgcC GGGCTCGCCA ttttcgccTA
651 CGTCAGCGGA CACTACCAAT TTTCCCAATA CCTCCAGCTT CCCTATGTCG
701 CCGGCGCGAA CGAAGTCGCT ATATTCTGCA CCGCCATGTG CGGCGCGTGC
751 CTCGGATTTT TGTGGTTCAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTGCGGCGC CTGGCATTGG GTGCGCGCTC CGGTaccGtt gCCGTcaTcg
851 tCCGCCAAGA ATTTGTcttc gtcattatGG GCGGTCTGTT cgtcgtagaa
901 gccgtgTCCG TTATGCTTCa tgtcggCTGG TACAAGAAAA Ccaaaaaacg
951 CATCTTcCTg acgGcaccga ttcacaccca ttaCCaactt cgatgCTGGa
1001 aagaaacgca agtcgctcgtc CGTTtCTGGA TTAtTAccat cgtcgtggtt
1051 tTgataggtt tGagtacctt caAAattcgc ggaaactatg ccgTCCGAAC
1101 ACCTTTcAGA CGGCATTTGA ACGCGCAATA A

```

This corresponds to the amino acid sequence <SEQ ID 284; ORF 088.ng>:

g088.pep

```

1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
51  KCGQAVRTDG PQTHLVKNGT PTMGGSLLT AITVSTLLWG NWANPYIWL
101 LGVLLATGAL GFYDDWRKVY YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
151 AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTGDLGL

```



201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC  
 251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLEFVVE  
 301 AVSVMLHVGW YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV  
 351 LIGLSTLKIR GNYAVRTPFR RHLNAQ\*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 285>:

m088.seq

```

1  ATGTTTTTAT GGCTCGCACA TTTCAGCAnC TGGTTAACCG GTCTGAATnn
51  nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
101 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
151 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
201 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
251 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
301 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
351 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
401 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
451 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
501 nnnnnnnnnn nnnGGCGTGG TCGGCTTTT GGTGTTGTCT TACCTGACCA
551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
601 GCGACCTTCC CCGTCGTCTC CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
651 TGCCAGCGGC CACTCACAAT TTGCCCAATA CCTGCAATTA CCTTACGTTG
701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTGGTGCA TTGGCATTGG GTGCCGCGCT CGGTACCGTC GCCGTATCG
851 TCCGCCAAGA GTTTGTCTCT GCATTATGG GCGGATTATT TGTCTAGAA
901 GCCGTATCCG TTATGCTTCA GGTGGCTGG TATAAGAAAA CCAAAAAACG
951 CATCTTCTG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
1001 AAGAAACCCA AGTCGTCGTC CGCTTTTGGG TTATTACCAT CGTCTTGGTG
1051 TTGATCGGTT TGAGTACCCT CAAAATCCG TGAACCTATG CCGTCTGAAC
1101 ATCTTTCAGA CGGCATTTGA ACGCGCAATA A

```

1 MFLWLAHFSN WLTGLNIFOY TTFRAVMAAL TALAFSLMEG PWTIRRLTAL  
 51 KCGQAVRTDG PQTHLVKNGT PTMGGSLLLT AITVSTLWLG NWANPYIWIL  
 101 LGVLLATGAL GFYDDWRKV YKDPNGVSAK FKMVWQSSVA VIAGLALFYL  
 151 AANSANNILI VPFFKQIALP LGVVGFVLVS YLTIVGTSNA VNLTDGLDGL  
 201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC  
 251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLEFVVE  
 301 AVSVMLQVGW YKKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV  
 351 LIGLSTLKIR XTYAVXTSFR RHLNAQ\*

This corresponds to the amino acid sequence <SEQ ID 286; ORF 088>:

m088.pep

```

1  MFLWLAHFSX WLTGLNXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
51  XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
101 XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
151 XXXXXXXXXXX XXXXXXXXXXX XGVVGFVLVS YLTIVGTSNA VNLTDGLDGL
201 ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLEFVVE
301 AVSVMLQVGW YKKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
351 LIGLSTLKIR XTYAVXTSFR RHLNAQ*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 088 shows 91.7% identity over a 205 aa overlap with a predicted ORF (ORF 088.ng)

from *N. gonorrhoeae*:

m088/g088

```

m088.pep                                10      20      30
                                GVVVGFVLVS YLTIVGTSNA VNLTDGLDGLA
                                |||
g088      TAGLALFYLAANSANNILIVPFFKQIALPLGVVGFVLVS YLTIVGTSNA VNLTDGLDGLA
                                150      160      170      180      190      200

```

282

	40	50	60	70	80	90
m088.pep	TFPVLVLAAGLAI FAYASGHSQFAQYLQLPYVAGANEVVI FCTAMCGACLGFLWFNAYPA					
	:					
g088	AFPVLVLAAGLAI FAYVSGHYQFSQYLQLPYVAGANEVAI FCTAMCGACLGFLWFNAYPA					
	21f	220	230	240	250	260
	100	110	120	130	140	150
m088.pep	QVFMGDVGALALGAALGT VAVIVRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
g088	QVFMGDVGALALGAALGT VAVIVRQEFVLVIMGGLFVVEAVSVMLHVGWYKTKKRIFLT					
	270	280	290	300	310	320
	160	170	180	190	200	
m088.pep	APIHHHYEQGWKETQVVVRFWIIITIVVLIGLSTLKIRXTYAVXTSFRRHNLNAQX					
	:					
g088	APIHHHYQLRCWKETQVVVRFWIIITIVVLIGLSTLKIRGNYAVRTPFRRHNLNAQX					
	330	340	350	360	370	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 287>:

```

a088.seq
1 ATGTTTTTAT GGCTCGCAC TTTCAGCAAC TGGTTAACCG GTCTGAATAT
51 TTTTCAATAC ACCACATTCC CGCGCGTACG GCGCGCGTTG ACCGCGTTGG
101 CGTTTTCCCT GATGTTCGGC CCGTGCACGA TACGCAGGCT GACCGCGCTCG
151 AAATGCGGGC AGGCAGTGCG TACCGACGGT CCGCAAACCC ACCTCGTCAA
201 AAACGGCAGC CCGACGATGG GCGGTTCCGT GATTCTGACG GCCATTACCG
251 TGTCACCCCT GTGTGTGGGG AACTGGGCCAA ACCCGTATAT CTGGATTCTC
301 TTGGGCGTAT TGCTCGGCAC GGGCGGCATC GGTTTTTACG ACAGCTGGCG
351 CAAAGTCGTC TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAATGG
401 TGTGGCAGTC AAGCGTTGCC ATTATCGCGC GTTTGGCATT GTTTTACCTT
451 GCCGCCAATT CCGGCCAACA TATTTTGATT GTCCCGTTCT TCAACCAATT
501 CGCCCTGCGC GTCTCGGCTG TCGGCTTTT GTGTTGPTCT TACGTACCA
551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
601 GCGACCTTCC CCGTCGTCTT CGTTGCGCGC GGCCTCGCCA TCTTCGCCTA
651 TGCCACGGCC CACTCACAAT TGTGCCAATA CCGTCAATTA CCTTACGTTT
701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGGCATGTG CGGCGCGTGC
751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTCGGTGCA TTGGCATTGG GTCCGCGCGT CGGTACCGTC GCCGTCATCG
851 TCCGCCAAGA TTTGTCTCTC GTCCATTATG GCGGATTATT TGTCTGATAG
901 GCGGTATCCG TTATGCTTCA GGTGCGGTGG TATAAGAAAA CCAAAAAAG
951 CATCTTCTCT ATGGCGCCCA TCCATACCA CTACGAACAA AAAGGCTGGA
1001 AAGAAACCCA AGTCGTGCTC CGTTTGTGTA TTATTACCAT CGTCTGGTG
1051 TTGATCGGTT TGAGTACCTT CAAATCCGCG TGAACCTATG CCGTCTGAAC
1101 ACCTTTTACA CGGCAATTGA ACCGCAATA A

```

This corresponds to the amino acid sequence <SEQ ID 288; ORF 088.a>:

a088.pep

1	MFLWLAHFSN	WLTGLNIFYQ	TFRAVMAAL	TALAFSLMFG	PWTIRRLTAL
51	KCGQAVRTDG	PQTHLVKNGT	PTMGQSLLT	AITVSTLLWG	NWANPYIWL
101	LGVLATGAL	GFYDDWRKVV	YKDPNGVSAK	FKMVWQSSVA	IIAGLALYIL
151	AANSANNILI	VPFFKQIALP	LGVVGFLVLS	YLTIVGTSNA	VNLTDLGLDL
201	ATFFVVLVAA	GLAIFAYASG	HSQFAQYLQL	PYVAGANEV	IFCTAMCGAC
251	LGFLWFNAYP	AQVFMGDVGA	LALGAALGT	AVIVRQEFVL	VIMGGLFVVE
301	AVSVMLQVWG	YKTKKKRIFL	MAPIHHHYEQ	KGWKETQVNV	RFWIITIVLV
351	LIGLSTLKIR	*TYAV*TPFR	RHLNAQ*		

**m088/a088** 99.5% identity over a 205 aa overlap

```

                150      160      170      180      190      200
m088.pep  XXXXXXXXXXXXXXXXXXXXXXXXXXXXGWWGFLVLSYLTIVGTSNAVNLTGDLGDLA
                |||||
a088      IAGLALFYLAANSANNILIVPFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTGDLGDLA

```

283

	150	160	170	180	190	200
	210	220	230	240	250	260
m088.pep	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVFCTAMCGACLGFLWFNAYPA					
a088	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVFCTAMCGACLGFLWFNAYPA					
	210	220	230	240	250	260
	270	280	290	300	310	320
m088.pep	QVFMGDVGLALGAALGTAVIVRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
a088	QVFMGDVGLALGAALGTAVIVRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
	270	280	290	300	310	320
	330	340	350	360	370	
m088.pep	APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRLNAQX					
a088	APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRLNAQX					
	330	340	350	360	370	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 289>:

```

g089.seq
1  ATGCCGCCCA  AAATCACGAA  GAGCGGGTTT  TGCAAACCGG  CAATCGCGGC
51  GGCGGTCGCG  CCGACATTCG  TGCCTTTGCT  GTCGTCGATG  AATACCACGC
101 CGTTTTTCTC  GCCGATTTT  TCCACACGGT  GCGGCAAGCC  TTGGAAGGTT
151 TTGACGTGTT  CCAGCAATGC  TTCGCGCGGC  AAACCGACGG  CCTCGCACAA
201 AGCCACGGCA  GCCATAACGT  TGGCGGCGTT  GTGCAAACCT  TGCAGCGGGA
251 TGTCTTGCGT  AGAAATCAAA  TCTTCATTGC  CTTGTTTTAA  ACAGCCCGTC
301 CCGCGTTCCA  ACCAAAAATC  GGCTTCGTGT  TCCAAGGAAA  ACCGTTTCAC
351 TTCACGCCCT  GCCCGTTTCA  TGGCGCGGCA  GAACACGTCG  TCCGCATTCA
401 AAACCTGCAC  TCCATCGCCA  CGGAAAATCT  CGGCTTTGGT  ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 290; ORF 089.ng>:

```

g089.pep
1  MPPKITKSGF  CKPAIAAAVA  PTFVPLLSSM  NTPPFSPPIF  STRCGKPWKV
51  LTCSSNASRG  KPTASHKATA  AITLAALCKP  CSGMSCVEIK  SSLPCFKQPV
101 PRSNOKSASC  SKENRFTSRP  ARPMARQNTS  SAFKTCTPSP  RKISALVCA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 291>:

```

m089.seq
1  ATGCCGCCCA  AAATCACKAw  GAGCGGATTT  TGCAAACCGG  CAATCGCGGC
51  GGCAGTCGCG  CCGACATTCG  TGCCTTTGCT  GTCGTCGATA  AACACCACGC
101 CGTTTTTCTC  GCCGATTTT  TCCACGCGGT  GCGGCAGGCC  TTGGAAGGTT
151 TTGACGTGTT  CGAGCAATGC  TTCGCGCGAC  AAACCGATGG  CCTCACACAA
201 AGCCACKGCA  GCCATGACGT  TAGCGGCGTT  GTGCAKACCT  TGCAACGGWA
251 TGTCTTGCGT  GACAATCAAA  TCTTCATTGC  CTTGTTTCAG  GCGGCCTGTC
301 TCGCGTTCCA  ACCAGAAATC  AGCTTCGTGT  TCCAACGAAA  ACCATTTTAC
351 CTCGCGCCCG  GCACGCTTCA  TCGCGCGGCA  GAACGCATCG  TCCGCATTCA
401 AAACCTGCAC  GCCGTCGCCA  CGGAAAATCT  TGGCTTTGGT  ATGCGCATAG

```

This corresponds to the amino acid sequence <SEQ ID 292; ORF 089>:

```

m089.pep
1  MPPKITXSGF  CKPAIAAAVA  PTFVPLLSSI  NTPPFSPPIF  STRCGRPWKV
51  LTCSSNASRD  KPMASHKATA  AMTLAALCXP  CNGMSCVTIK  SSLPCFRRPV
101 SRSNOKSASC  SNENHFTSRP  ARFIARQNAS  SAFKTCTPSP  RKILALVCA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 089 shows 88.6% identity over a 149 aa overlap with a predicted ORF (ORF 089.ng) from *N. gonorrhoeae*:

m089/g089

284

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPAIAAAVAPTFVPLLSINTT	PFFSPIFSTRCGRPWKVLTCSSNASRD				
g089	MPPKITXSGFCKPAIAAAVAPTFVPLLSMNTT	PFFSPIFSTRCGRPWKVLTCSSNASRG				
	10	20	30	40	50	60
	70	80	90	100	110	120
m089.pep	KPMASHKATAAMTLAALCXPNGMSCVTIKSSLPCFRRPV	SRSNQKSASCSNENHFTSRP				
g089	KPTASHKATAAITLAALCKPCSGMSCVEIKSSLPCFKQPV	PRSNQKSASCSKENRFTSRP				
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARQNASSAFKTCTPSPRKILALVCAX					
g089	ARFMARQNTSSAFKTCTPSPRKISALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 293>:

a089.seq

```

1  ATGCCGCCTA AAATCACGAA GAGCGGATTT TGCAAACCGG CAATCGCGGC
51  GCGCGTCCGA CCGACGTTTC TGCCTTTGCT GTCGTCGATG AACACCACGC
101 CATTTTTCTC GCCGATTTT TCCACGCGGT GCGGCAGGCC TTGAAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGGC AAACCGACGG CTTCGCACAA
201 GGCAACGGCA GCCATCACGT TAGTGGCGTT GTGCAAGCCT TGCAGCGGAA
251 TATCTTGCGT GGCAATCAAA TCTTCATTGC CTTGTTTCAG GCGACCTGTC
301 TCACGTTCCA ACCAAAAATC GGCTTCGTAT TCCAACGAAA ACCATTTTCA
351 CTCGCGCCCG GCGCGCTTCA TCGCACGACA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC ACCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 294; ORF 089.a>:

a089.pep

```

1  MPPKITKSGF CKPAIAAAVA PTFVPLLSM NTTPFFSPIF STRCGRP*KV
51  LTCSSNASRG KPTASHKATA AITLVALCKP CSGISCVAIK SSLPCFRRPV
101 SRSNQKSASY SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

m089/a089 91.9% identity over a 149 aa overlap

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPAIAAAVAPTFVPLLSINTT	PFFSPIFSTRCGRPWKVLTCSSNASRD				
a089	MPPKITKSGFCKPAIAAAVAPTFVPLLSMNTT	PFFSPIFSTRCGRPWKVLTCSSNASRG				
	10	20	30	40	50	60
	70	80	90	100	110	120
m089.pep	KPMASHKATAAMTLAALCXPNGMSCVTIKSSLPCFRRPV	SRSNQKSASCSNENHFTSRP				
a089	KPTASHKATAAITLVALCKPCSGISCVAIKSSLPCFRRPV	SRSNQKSASYSNENHFTSRP				
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARQNASSAFKTCTPSPRKILALVCAX					
a089	ARFIARQNASSAFKTCTPSPRKILALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 295>:

g090.seq

```

1  ATGCGCGTAG TCGAGCAAAT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA

```

285

```

51  TGTTTCATCAC CGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
101 TGGAAGCTGG AAAGCTCcca CACCCACACG TCCGCCTTTT TGCCTTCGcg
151 ctgCAATtct gcctccaaga cgggcgtacc gatATTGCCC GCAATGAcgg
201 tatccagccc gcacttgatg CAGAGatagc ggaccaggct ggttaccgTG
251 GTTttgccgt tgctgCcggt aatcgCaatc accttg:cgC CGCGGCGGtt
301 cAaatGTCC gccagCAATt ggATGTCGCC TAGCACGCGC .ccgccgTTT
351 TGCTtga

```

This corresponds to the amino acid sequence <SEQ ID 296; ORF 090.ng>:

```

g090.pep
1  MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFA
51  LQFCLQDGRt DIARNDGIQp ALDAEIAQDA CYFGFAVAAG NRNLVAAAV
101 HNVRRQLDVA XHAYRRFA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 297>:

```

m090.seq
1  ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGA
51  TGTTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
101 TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
151 CTGCCATTCC GCCTCCAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
201 TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
251 GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT .CCGCCGTTT
351 TGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 298; ORF 090>:

```

m090.pep
1  MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
51  LPFRLQNRRA DIARDNGIQp ALDTEIAQDA RYRGFAVAAG NRNYLVVPAV
101 HNVRRQFDVA QHAXRRFA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 090 shows 83.9% identity over a 118 aa overlap with a predicted ORF (ORF 090.ng)

from *N. gonorrhoeae*:

```

m090/g090
10      20      30      40      50      60
m090.pep MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFALPFR LQNRRA
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
g090      MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFALQFCLQDGRt
10      20      30      40      50      60

70      80      90      100     110     119
m090.pep DIARDNGIQPALDTEIAQARYRGFAVAAGNRNYLVVPAVHNVRRQFDVAQHAXRRFAX
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
g090      DIARNDGIQPALDAEIAQAGYRGFAVAAGNRNLVAAAVHNVRRQLDVAXHAXRRFAX
70      80      90      100     110

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 299>:

```

a090.seq
1  ATGCGCGTAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGA
51  TGTTTCAGCAC TGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
101 TGGAAACTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
151 CTGCAATTCC GCCTCCAAA CCGGCGCGCC GATATTGCCC GCGATAACGG
201 TATCCAGCCC ACATTGATG CAGAGATAGC CGACCAGGCT CGTTACCGTG
251 GTTTTGCCGT TGCTGCCGGT AATCGCAATC ACCTTGTCGC CGCGGCGGTT
301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT C.CGCCGTTT
351 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 300; ORF 090.a>:

286

**a090.pep**

```

1 MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQ HPHVRLFAFA
51 LQFRLQNRRA DIARDNGIQP TLDAEIADQA RYRGFAVAAG NRNHLVAAAV
101 HNVROQFDVA QHAXRRFA*

```

**m09/a090** 91.5% identity over a 117 aa overlap

```

              10      20      30      40      50      60
m090.pep    MRIVEQVVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA
              ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a090         MRVVEQVVVAVEMVFGNVQHCRRSRAQAFGVFQLETGKLQHPHVRLFAFALQFRLQNRRA
              10      20      30      40      50      60

              70      80      90     100     110     119
m090.pep    DIARDNGIQPALDTEIADQARYRGFAVAAGNRNRYLVVPAVHNVRQQFDVAQHAXRRFAX
              ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a090         DIARDNGIQPTLDAEIADQARYRGFAVAAGNRNHLVAAAVHNVRQQFDVAQHAXRRFAX
              70      80      90     100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae***g090-1.seq** This sequence contains multiple stop codons (not shown)

This corresponds to the amino acid sequence &lt; ORF 090-1.ng&gt;:

**g090-1.pep** (not shown)The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 2>:

```

m090-1.seq
1 ATGACGGCGT TTGCATTCA GACGGCATCA CAAAGCCTTA AACGCTTCGA
51 TAAACACTTC CGAACGGTGC GCGTAGCCTT TGAACATATC AAAGCTCGCG
101 CAGGCGGGGC TGAGCAACAC AATATCGCCT GCTTCGGCTT GGGCAT*TGC
151 CGTCTGAACG GCTTCTCCCA AAGTGGCGCA GTCGGTCATA TTCAAGCCGC
201 AGCCGTCCAA ATCGCGCGCG ATTTGCGGCG CATCGACACC AATCAAGAAC
251 ACGCCTTTTG CTTGCTTAC CAGTGCATCG CGCAGGGGCG TGAAGTCCTG
301 CCCTTTACCC ATGCCGCCCA AAATCACGAA GAGCGGATTT TGCAAAACCGG
351 CAATCGCGGC GGCACTCGCG CCGACATTCT TGCCTTTGCT GTCGTCGATA
401 AACACCACGC CGTTTTTCTC GCGGATTTTT TCCACGCGGT GCGGCAGGCC
451 TTGGAAGGTT TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG
501 CCTCACACAA AGCCACGGCA GCCATGACGT TAGCGGCGTT GTGCAGACCT
551 TGCAACGGAA TGTCTTGCCT GACAATCAAA TCTTCATTGC CTTGTTTICAG
601 GCGGCCTGTC TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA
651 ACCATTTTAC CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG
701 TCCGCATTCA AAACCTGCAC GCCGTGCGCA CGGAAAATCT TGGCTTTGGT
751 ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
801 TGTTTCAGCA CGTCGCGGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAST
851 TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
901 CTGCCATTCC GCCTCCAAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
951 TATCCAGCCC GCACCTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
1001 GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
1051 CACAATGTCC GCCAGCAATT CGATGTGCC CAACACGCGT CCGCCGTTTT
1101 GCTTGAACGC CTCAATATCC GGCTGCCGCT CGCTGATGCC GGGACTGAGA
1151 GCCAGAATAT CGAAACCGTT GTCCAGCGCA TCTTTCAGAC GGCCCGTGTA
1201 AAACACCAAC CGGTCAAACA TCTTACCGAT TTGCGACACG CGTTCGGCT
1251 TCAGCTCCGC ATCATAACGA GCAACCTCCG CGCCGTTTTT GCGCAGGTAG
1301 GCAATCATGG AAATACCGGT ACCGCCGAGT CCGGCGACGA GGATTTTTTT
1351 GTTTTGAAAA GTCATTTTGG TTTGTCTTAA

```

This corresponds to the amino acid sequence &lt;SEQ ID 3; ORF 090-1&gt;:

```

m090-1.pep
1 MTAFAPQTAS QSLKRFDKHF RTVRVAFEHI KARAGGAEQH NIACFGLGIC
51 RLNGFSQSGA VGHIAAAVQ IAADLRRIDT NOEHAFCLAY QCIAQGREV
101 PFTHAQNHE ERIQTGNRG GSRADIRAF VVDKHHAVFL ADFFHAVRQA
151 LEGFDVFEQC FARQTDGLTO SHGSHDVSGV VQTLQRNVLR DNQIFIALFQ
201 AACLAFAQEI SFVFORKPFY LAPGTLHRAA ERIVRIQNLH AVATENLGFG
251 MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
301 LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNRYLVVPAV
351 HNVROQFDVA QHASAVLLER LNIRLPLADA GTESQNIETV VQRIQFQARV
401 KHQPVKHLTD LRHAFRLQLR IIRSNLRAVF AQVGNHGHTR TAESGDEDF
451 VLKSHFGLS*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 303>:

```
g091.seq
1  ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTTTGA
51  AAGTCATTTT GGTTTGTGCC TAAACAAAT CATATTGGGC AGGAGACGTC
101 CGCCCTTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCCG ATTAATAACC
151 CGCCTTCAGG CGTTGGTCAT TGTCGCAGCT GTTTGGTCT CCGTTTGAC
201 AAGCCTTGCC AAGCCATTGT TGAGCGAGCG CAAGGCTTTG GCGCACGCCG
251 CGTCCATCGT AATACATCAA GCCCAAATTG TATTGGGCTT GGGCATCCCC
301 TTGTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 304; ORF 091.ng>:

```
g091.pep
1  MEIPVPPSPA TRIFLFESHF GFVLKQIILG RRRPPLPKPL SDGIASRLIT
51  RLQALVIVAA VLVSVLTSLA KPLLSEKVL AHAASIVIHQ AQIVLGLGIP
101  LF*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 305>:

```
m091.seq
1  ATGGAAATAC CCGTACCGCC GAGTCCGGCG ACGAGGATTT TTTTGTTTGA
51  AAAGTCATTT TGGTTGTGCC TAAACAAAT CATATTGAGC AGGAGATGTC
101 CGCCCTTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
151 CGCCTTCAGG CGTTGGTCAT TGTCGCAGCC GTCTTGGTCT CCGTTTGAC
201 AAGCCTTGCC AAACCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
251 CGTCTTTCGG CATACATCAC GCCCAAATTG TTTTGGGCTT GGGCTACCCC
301 CTGCGC...
```

This corresponds to the amino acid sequence <SEQ ID 306; ORF 091>:

```
m091.pep
1  MEIPVPPSPA TRIFLFEKSF WFLVKQIILS RRCPPPLPKPL SDGIASCSIT
51  RLQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLGLGYP
101  LR
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 091 shows 84.2% identity over a 101 aa overlap with a predicted ORF (ORF 091.ng) from *N. gonorrhoeae*:

m091/g091

	10	20	30	40	50	60
m091.pep	MEIPVPPSPATRIFLFEKSFVFLVKQIILSRRCPPPLPKPLSDGIASCSITRLQALVIVAA					
	:     :     :     :     :     :					
g091	MEIPVPPSPATRIFLFESHFGFVLVKQIILGRRPPLPKPLSDGIASRLITRLQALVIVAA					
	10	20	30	40	50	60
	70	80	90	100		
m091.pep	VLVSVLTSLAKPFLCKGAVLAHAASFGIHHQAQIVLGLGYPLR					
	:     :     :     :     :     :					
g091	VLVSVLTSLAKPLLSEKVLHAHAASIVIHQAQIVLGLGIPLFX					
	70	80	90	100		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 307>:

```
a091.seq
1  ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTTTTG
51  GAAATCATTT TGGTTGTGCC TAAACAAAT CATATTGAGC AGGGGATGTC
101 TGATCCTGCT CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
151 CGCTTTCAGG CGTTGGTCAT TGTCGCAGCT GTCTTGGTAT CCGTTTGAC
201 AAGCCTTGCC AAGCCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
251 CGTCTTTCGG CATACATCAC GCCCAAATTG TTTTGGGC
```

This corresponds to the amino acid sequence <SEQ ID 308; ORF 091.a>:

a091.pep

```

1  MEIPVPPSPA TRIFLWKSF WFVLKQIILS RGCLILLKPL SDGIASCSIT
51  RFQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLG

```

m091/a091 93.8% identity over a 96 aa overlap

```

              10      20      30      40      50      60
m091.pep  MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPLPKPLSDGIASCSITRLQALVIVAA
          |||
a091      MEIPVPPSPATRIFLFEKSFWFVLKQIILSRGCLILLKPLSDGIASCSITRFQALVIVAA
              10      20      30      40      50      60

              70      80      90      100
m091.pep  VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLGLGYPLR
          |||
a091      VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLG
              70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 309>:

g092.seq

```

1  ATGTTTTTAA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGTGCGC
51  AACCGGTCAG ACCTTTAAAA TAACGCCTTT ACGCACTAAA AACCAACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTAA GCAACATCCA TTTGTGCGGT
151 ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTGGG
201 CTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251 TGAGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
301 AACGGTgcgg ATGTCGTCGT TGCCTCTACC GCCGTCAGA AAGAAaatcC
351 CGAAGTtgtc gcTGC GTTGG AGCGGCAAAAT TCCCGTTATT CCGCGCGCCT
401 TGATGCTGGC AGAGCTGATG CGCTTCCGTG ACgcatcgcc cattgcccgt
451 ACGCACGGCA AAACACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTC GACCCCACTT TCGTTATCGG CGGCAAACTC AACGCCGCGC
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAATCCGATG CCTCTTTCCT ACATCTGACC CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTGTTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAACTT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCAAAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGc gtggcGCTgg
1001 aagtcGgCGC ATcggttgAA GCGAtcCAA AaggCTTGCT CGGCTTTGAA
1051 GGCGTCGGCC GCCGCTTCCA AAAATAcggc gacatCAagt tgccaaacgg
1101 cggGaccgCT TTgctGGTGG ACGATTAcgg ACACCACCCC GTCGAAATGG
1151 CGGcaaccct tgccgcTGCA CGCGGCGCGT ATCCGAAAAA acgtTTGGTG
1201 CtcgCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTACTCAATA CCGTTGatgC GCTGGTACTG ACCGAAAGTTT
1301 AtgcccgcgG CGAAGAGCCG GTTGCCGCGC CGGactcCCG CGCCTTGGCG
1351 CGTGCTATCC GCGTATTGGG CAAACTTGAG CCGATTTACT GCGAAAatgt
1401 cgcgcACCTG CCGCAAATGC TGATGAATGT TTTACAGGAT Gcgcatgttg
1451 tggtgAATAT GggTgcggga agcatcaacc gcgttccttc cgcgctgttg
1501 gaattgtcga AACAGAtttg A

```

This corresponds to the amino acid sequence <SEQ ID 310; ORF 092.ng>:

g092.pep

```

1  MFFISIRYIF VRKLWCANGQ TFKITPLRTK NQPERNIMMK NRVSNIHFVG
51  IGGVGMMSGIA EVLHNLGPKV SGSDQARNAA TEHLSSLGIQ VYPGHAEHV
101 NGADV VAST AVKKENPEVV AALERQIPVI PRALMLAELM RFRDGLAIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK

```



289

```

251 AFLCVDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEV GASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP VAAADSRALA
451 RAIRVLGKLE PIYCENVADL PQMLMNVLQD GDVVLNMGAG SINRVPSALL
501 ELSKQI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 311>:

```

m092.seq
1 ATGTTT TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
51 AAACGGTCAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTTA CTAACATCCA TTTTGTCCGT
151 ATCGGCGGCG TCGTCATGAG CGGCATCGCC GAGTCTTGC ACAATTTGGG
201 CTTTAAAGTT TCCGCTTCGG ATCAGGCGCG AAAAGCCGCT ACCGAGCATT
251 TGGGCAGCCT GGCATTCAA GTTTATCCCG GCCATACCGC CGAACACGTT
301 AACGGTGCAG ATGTCGTCGT TACCTCTACC GCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GTCGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTCGCGGC
451 ACGCACGGCA AAACCACGAC CACGAGCCTG ACCGCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCGAG
551 GCACTAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGT GTATTGACAG CGAACACGTC CGCGCGATT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCGCAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGCGCG ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GCGTCGCGCC GCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 CGGGACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CGGCGACCC TCCGCGCGCA CGCGCGCGT ATCTGGAAAA ACGTTTGGTA
1201 CTCGCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAA GTCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CCGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTGGG CAAACTCGAG CCGATTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTGAACGT TTTGACGAG GCGACATCG
1451 TGTGAATAT GGGCGCGGGA AGCATCAACC GCGTCCCGC CGCGCTGCTG
1501 GCATTGTCGA AACAGATTTG A

```

This corresponds to the amino acid sequence <SEQ ID 312; ORF 092>:

```

m092.pep
1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVNTNIHFVG
51 IGGVMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101 NGADVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEV GASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYLEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
501 ALSKQI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 092 shows 96.6% identity over a 506 aa overlap with a predicted ORF (ORF 092.ng) from *N. gonorrhoeae*:

m092/g092

```

10      20      30      40      50      60
m092.pep MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVMSGIA

```

a092.seq

1	ATGTTTTTTT	TTTCAATCCG	CTATATATTT	GTCAGAAAAAC	TATGGCGCGC
51	AAACCGCTCAG	CCCTTTAAAA	TAACGCGCTT	ACGATCTCGAA	ATATCCAGCG
101	AACGCAATCA	TATGATGAAA	CCGCGAGTGA	CAACATCCCA	TTTTTCGGT
151	ATCGGCGCGC	TCGGCATGAG	CGGTATCGCC	GAAGTCTTGC	ACAATTTGGG
201	TTTTTAAAGT	TCGGTTTCGG	ATCAGGCGCG	GGATGCGGCT	ACCGGACAT
251	TGGGCGACCT	GGGCATCTAA	GTTTATCCCG	CCCATACCCG	AGAACAGGTT

291

```

301 AACGGTGGCG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGT GTTGAGAAGC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACGGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCAAAA TGAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CCGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GGTGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 TGGAACCGCG CTCTTGGTGG ACGACTACGG ACACCAACCC GTCGAAATGG
1151 CCGCGACCCT TTCCGCGCGC ACGCGCGCGT ATCCGGAATA ACGTTTGGA
1201 CTCGCCCTTC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCGCG TGAAGAGCCG ATTGCCGCG CTGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTGGG CAAACTCGAG CCGATTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTGAACGT TTTGCAGGAC GGCGACATCG
1451 TGTGAATAT CCGTGGCGGA AGCATCAACC GCGTCCCGC CGCGCTGCTG
1501 GAATTGTCGA AACAGATTG A

```

This corresponds to the amino acid sequence <SEQ ID 314; ORF 092.a>:

#### a092.pep

```

1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIEHVG
51 IGVGMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101 NGADVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGLAIAIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMFFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKTIVHV
301 QMGHEQGSF EVVLNMPGRH NVLNALAAIG VALEV GASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPPNGTA LLVDDYGHHP VEMAATLSAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRLA
451 RAIRVLGKLE PIYCENVADL PEMLNVLDQ GDIVLNMGAG SINRVPAALL
501 ELSKQI*

```

m092/a092 99.4% identity over a 506 aa overlap

	10	20	30	40	50	60
m092.pep	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFGVIGGVGMSGIA					
a092	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFGVIGGVGMSGIA					
	10	20	30	40	50	60
m092.pep	EVLHNLGFKVSGSDQARNAAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKENPEVV					
a092	EVLHNLGFKVSGSDQARNAAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKENPEVV					
	70	80	90	100	110	120
m092.pep	EVLHNLGFKVSGSDQARNAAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKENPEVV					
a092	EVLHNLGFKVSGSDQARNAAATEHLGSLGIQVYPGHTAEHVNGADVVTSTAVKKENPEVV					
	130	140	150	160	170	180
m092.pep	AALEQQIPVIPRALMLAELMRFRDGLAIAIAGTHGKTTTSLTASILGAAGLDPTFVIGGKL					
a092	AALEQQIPVIPRALMLAELMRFRDGLAIAIAGTHGKTTTSLTASILGAAGLDPTFVIGGKL					
	130	140	150	160	170	180
m092.pep	NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					
a092	NAAGTNARLGKGEYIVAEADESDASFLHLTPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					

292

	190	200	210	220	230	240
	250	260	270	280	290	300
m092.pep	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKTVHV					
a092	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKTVHV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m092.pep	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEVGASVEAIQKGLGFEGVGRRFQKYG					
a092	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEVGASVEAIQKGLGFEGVGRRFQKYG					
	310	320	330	340	350	360
	370	380	390	400	410	420
m092.pep	DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLAFAQPHRYTRTRDLFEDFTK					
a092	DIKLPNGGTALLVDDYGHHPVEMAATLSAARGAYPEKRLVLAFAQPHRYTRTRDLFEDFTK					
	370	380	390	400	410	420
	430	440	450	460	470	480
m092.pep	VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
a092	VLNTVDALVLTEVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
	430	440	450	460	470	480
	490	500				
m092.pep	GDIVLNMGAGSINRVPAALLALSQIX					
a092	GDIVLNMGAGSINRVPAALLELSQIX					
	490	500				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 315>:

```

g093.seq
1  aTGCAGAATt ttgGCAAAGT ggccgtATTG ATGGGtggtT TTTCCAGCGA
51  ACGAGAAatc tcgcTGGACA GcGTACCGC CATTTTGAAC GCCTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGACC CTAAGGAAAC GCCGTATATC
151 GAACTGAAGG AGCGGGGCTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAAGTGT TGGGCAATTC
251 CCTATACCGG CAGCGGTGTC GCCGCCCTCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAc TGATTTGGCA GGCATTGGGA TTACCCGTTT CCGAGTTCGC
351 CGTACTGTAC GATGATACCG ATTTTCGATGC CGTCGAAGAA AAATTGGGTC
401 TGCCGATGTT TGTGAAGCCG GCGGCCGAAG GCAGCAGCgt cggcgtggtA
451 aAAGTCAAAG AAAaaggccg TCTGAAAAGC GTTtracgaag aatTGAAaCA
501 CCTTcagggg cgaAAtcatt gccgAacgTT TTATCGGCGG CCGCGAATAT
551 TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCGGCATAC ACATCATCCC
601 CGCAACCGAG TTTTACGAct acgaagccaa GtaciaaCCGA GACGAcacca
651 tttAACAATG TCCTTCGGAA GATTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTTCGCG CGCACAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGAAAA
801 TCAACACCTT GCCCGGTATG ACCGGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 316; ORF 093.ng>:

```

g093.pep
1  MQNFGKVAVL MGFSSEREI SLDSGTAILN ALYSKGIDAY AFDPKETPLS
51  ELKERGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLY DDTDFDAVEE KLGLPMFVKP AAESSVGVV
151 KVKEKGRLLS VYEELKHLQG RNHCRTFYRR RRIFLPRPER QRAARHTHP
201 RNRVLRRLSQ VQPRRHLSM SFGFRDRSRR KPDARTGGSR RTGNRCGRLE
251 ARRFQRYRR QTLVGNQHP ARYDRP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 317>:

```

m093.seq

```

m093.ppt

1	MONFGKVAVL	MGGFSSEREI	SLDSGTAILN	ALKSKGIDAY	AFDPKETPLS
51	ELKKAQGFQTA	FNILHGYTERI	DGAVQGALEL	LGLPYTGSVG	AASAI GMDKY
101	RCKLIQWALG	LPVPEFAVLH	DDTFDVAEE	LKGI PMFVKP	AAEGSSGVGV
151	KVKKGKRLKS	VYEELKHLOX	RNHCRTFYRR	RRIFLPRPER	QRAARHTHHS
201	RNRVLRLRSQ	QVPRRHHLMS	SFGRFDRSR	KPDARTGGSR	RAGNRCCRRLR
251	ARRFPORYRY	OTLSVGNCHP	ARYDEP*		

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 093 shows 96.7% identity over a 276 aa overlap with a predicted ORF (ORF 093.ng) from *N. gonorrhoeae*:

m093/g093

		10	20	30	40	50	60
m093 . pep		MQNFGKVAVL	MGGFSSEREIS	LDSGTAILN	ALKSKGIDAY	AFDPKETPL	SELKAQGFQTA
g093		MQNFGKVAVL	MGGFSSEREIS	LDSGTAILN	ALKSKGIDAY	AFDPKETPL	SELKERGFQTA
		10	20	30	40	50	60
		70	80	90	100	110	120
m093 . pep		FNILHGT	YGXDGA	VQGALELL	GIPTGSG	VAAISAIG	MDKYRCKLI
g093		FNILHGT	YGEDGA	VQGALELL	GIPTGSG	VAAISAIG	MDKYRCKLI
		70	80	90	100	110	120
		130	140	150	160	170	180
m093 . pep		DDTDFDA	VEEKLGL	PMFVKPA	AEGSSVG	VVKVKGK	GRLSVYEE
g093		DDTDFDA	VEEKLGL	PMFVKPA	AEGSSVG	VVKVKEK	GRLSVYEE
		130	140	150	160	170	180
		190	200	210	220	230	240
m093 . pep		RRIFLPR	PERQRAA	RHTHHSR	NRVLRLR	SQVQPRR	HHLSMSF
g093		RRIFLPR	PERQRAA	RHTHHPN	RVLRSLR	SQVQPRR	HHLSMSF
		190	200	210	220	230	240
		250	260	270			
m093 . pep		RAGNRCG	RLRARRF	PQRYRRQ	TL	SVGNQH	PARYDEPX
g093		RTGNRCG	RLRARRF	PQRYRRQ	TL	SVGNQH	PARYDRPX
		250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 319>:

**a093.seq**

```

1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CCAAGGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGCGCAA GACGGGGCTG TTCAGGGTGC ATTGGAAC TGCGGCATTG
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTGATGC CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAAACCG GCGGCCGAG GCAGCAGCT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTACGAAG AATTGAAACA
501 CTTTCAGGG. CGAAATCATT GCCGAACGGT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCTG TGTGAAACCG CAAAGGCCTG CCCGGCATA ACATCATCCC
601 CGCGACCGAG TTTTATGACT ACGAAGCCAA GTACAACCGC AACGACACCA
651 TTTATCAATG TCCTTCGGAA GATCTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTGCGCG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAG
801 TCAACACCCT GCCCGGTATG ACCGGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 320; ORF 093.a>:

**a093.pep**

```

1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
151 KVKGKGR LKS VYEELKHFQX RNHCRTVYRR RRIFLPCVER QRPARTHH
201 RDRVL* LRSQ VQQRHHLMS SFGSRDSRRR KPDARTGGSR RAGNRCGRLR
251 ARRFQRYRR QTLVGNQHP ARYDRP*

```

**m093/a093** 95.7% identity over a 276 aa overlap

	10	20	30	40	50	60
m093.pep	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS ELKAQGFQTA					
a093	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS ELKAQGFQTA					
	10	20	30	40	50	60
m093.pep	FNILHGTYGXD GAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
a093	FNILHGTYGED GAVQGALELLGIPYTGSGVAASAIGMDKYRCKLIWQALGLPVPEFAVLH					
	70	80	90	100	110	120
m093.pep	DDTDFDAVEE KLGLPMFVKPAAEGSSVGVV KVKGKGR LKSVYEELKHLQXRNHCRTFYRR					
a093	DDTDFDAVEE KLGLPMFVKPAAEGSSVGVV KVKGKGR LKSVYEELKHFQXRNHCRTVYRR					
	130	140	150	160	170	180
m093.pep	RRIFLPRPERQRAARHTHHSRNRVLR LRSQVQPRRHLSMSFGRFDRSRRKPDARTGGSR					
a093	RRIFLPCVERQRPARTHHPRDRVLX LRSQVQQRHHLMSFGRSDRSRRKPDARTGGSR					
	190	200	210	220	230	240
m093.pep	RAGNRCGRLRARRFPQRYRRQTLVGNQH PARYDEPX					
a093	RAGNRCGRLRARRFPQRYRRQTLVGNQH PARYDRPX					
	250	260	270			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 321>:

```
g094.seq
1  ATGTATTTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
51  GCCGCCGATA ACGAAAGTGG GGTGAGTCC TGCCGCGCCG AGGATGGAGG
101 CCGTCAGGCT GGTGGTCGTG GTTTGCCGT GCGTAccggc aatggcgatg
151 cCGTCACGGA AGCGCATCAG CTCTGCCAGC ATCAAGGCGC GCGGAATAAC
201 GGGAAATTGC CGTCCAACG CAgcgacaAC TTCGGgattT TCTTTCTTGA
251 CGGCGGTAGA GGCAACGACG ACATccgcAC CGTTAACGTG TTCTGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 322; ORF 094.ng>:

```
g094.pep
1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPVCPAMAM
51  PSRKRISSAS IKARGITGIC RSNAATTSGF SFLTAVEATT TSAPLTC SAV
101 WPG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 323>:

```
m094.seq
1  ATGTATTTCGC CTTTGCCCAA GCGGGCGTTA GTGCCTGCGG CGTTGAGTTT
51  GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
101 CCGTCAGGCT GGTGGTCGTG GTTTGCCGT GCGTGCCGGC AATGGCGATG
151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
201 GGGAAATTGC TGCTCCAACG CAGCGACAAC TTCGGGATT TCTTTTTTGA
251 CGGCGGTAGA GGTAAACGACG ACATCCGCAC CGTTAACGTG TTCGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 324; ORF 094>:

```
m094.pep
1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPVCPAMAM
51  PSKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTC SAV
101 WPG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 094 shows 95.1% identity over a 103 aa overlap with a predicted ORF (ORF 094.ng) from *N. gonorrhoeae*:

```
m094/g094

          10      20      30      40      50      60
m094.pep  MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVV VLPVCPAMAMP SRKRINSAN
          |||||
g094      MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVV VLPVCPAMAMP SRKRISAS
          10      20      30      40      50      60

          70      80      90     100
m094.pep  IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTC SAVWPGX
          |||||
g094      IKARGITGICRSNAATTSGFSFLTAVEATTTSAPLTC SAVWPGX
          70      80      90     100
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 325>:

```
a094.seq
1  ATGTATTTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
51  GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
101 CCGTCAGGCT GGTGGTCGTG GTTTGCCGT GCGTGCCGGC AATGGCGATG
151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
201 GGGAAATTGC TGCTCCAACG CAGCGACAAC TTCGGGATT TCTTTTTTGA
251 CGGCGGTAGA GGTAAACGACG ACATCCGCAC CGTTAACGTG TTCTGCGGTA
301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 326; ORF 094.a>:

a094.pep  
 1 MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM  
 51 PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV  
 101 WPG\*

m094/a094 100.0% identity over a 103 aa overlap

	10	20	30	40	50	60
m094.pep	MYSPLPKRALVPAALSLPPI	TKVGSSPAAPRMEAVRLVVV	VLPCVPAMAMESRKRINSAN			
a094	MYSPLPKRALVPAALSLPPI	TKVGSSPAAPRMEAVRLVVV	VLPCVPAMAMESRKRINSAN			
	10	20	30	40	50	60

	70	80	90	100
m094.pep	IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX			
a094	IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX			
	70	80	90	100

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 327>:

g095.seq  
 1 ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTGCG GCGCCGACGT  
 51 TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA  
 101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC  
 151 AACACACAAA AAGGCTTTGC CGTAGAAGGG CATAACGGTGG ATGAAATCGA  
 201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG  
 251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGGGTCA GTGTAGGAAA  
 301 GAGGCATCGG ATCGTCGGCT TCGGCAACGA TGTATTCCGC TTTGCCCAAG  
 351 CGGGCGTTGG TGCCTGCGGC GTTGA

This corresponds to the amino acid sequence <SEQ ID 328; ORF 095.ng>:

g095.pep  
 1 MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV  
 51 NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRGQCRK  
 101 EASDRRLRQR CIRLCPSGRW CLRR\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 329>:

m095.seq  
 1 ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTGCG GCGCCGACGT  
 51 TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA  
 101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC  
 151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATAACGGTGG ATGAAATCGA  
 201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG  
 251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG  
 301 GATGCATCGG ACCGTCGGCT TCGGCAACGA TGTATTCCGC TTTGCCCAAG  
 351 CGGGCGTTAG TGCCTGCGGC GTTGA

This corresponds to the amino acid sequence <SEQ ID 330; ORF 095>:

m095.pep  
 1 MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV  
 51 NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRGQCRK  
 101 DASDRRLRQR CIRLCPSGRX CLRR\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 095 shows 97.6% identity over a 124 aa overlap with a predicted ORF (ORF 095.ng) from *N. gonorrhoeae*:

m095/g095

	10	20	30	40	50	60
m095.pep	MSFHLNMDGEFHLRADVFDV	GGVDVGGIVQTVSSIRFAHF	GQNRADVFAVNTQKGFAVEG			



297

```

g095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFVEG
           10      20      30      40      50      60

           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
g095      HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKEASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
g095      CLRRX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 331>:

```

a095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATA CGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGCTTC TCAACACTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301 GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351 CGGGCGTTGG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 332; ORF 095.a>:

```

a095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF QNRADVFAV
51  NTQKGFVEG HTVDEIDKRL MQLLNTVPVG IHMFVDIGN DGHNRQCRCR
101 DASDRRLRQR CIRLCPSGRW CLRR*

```

m095/a095 96.0% identity in 124 aa overlap

```

           10      20      30      40      50      60
m095.pep  MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFVEG
           |||||
a095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFVEG
           10      20      30      40      50      60

           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
a095      HTVDEIDKRLMQLLNTVPVG IHMFVDIGN DGHNRQCRCR KDASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
a095      CLRRX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 333>:

```

g096.seq
1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGTGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGAtaaat ggtgTCGTCT CGGttgtaCt

```

298

```

201  tggcttcgta  gTCTTAAAC  TCGGTTGCGG  GGATGATGTG  TATGCCGGGC
251  AGCCCTTTGC  CGTTCAGGAC  GGGGCAGGAA  TATTCGCCGC  CGCCGATAAA
301  AcgtTcggca  atgaTTcgc  ccctgAAGGT  GttTCAattc  ttcgtaAACG
351  CTTTTCAGAc  ggccttTTTC  TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 334; ORF 096.ng>:

**g096.pep**

```

1  MAGHTGQGVDFQQIEFAVGI FEEIDAHAAF RTDCLCAANR QFAHQAFFGF
51  GQIFRRTLIN GVVSVVLGFV VVKLGCGDDV YAGQPFVAVQD GAGIFAAADK
101 TFGNDFAPEG VSILRKRFSD GLFL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 335>:

**m096.seq**

```

1  ATGGCTCGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGCGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTG CGGTTGTACT
201 TGGCTTCGTA GTCGTAAAAC TCGGTTGCGG GAATGATGTG TATGCCGGGC
251 AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 ACCTTCGGCA ATGATTTCGC CC.TGAAGGT GTTCAATTTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 336; ORF 096>:

**m096.pep**

```

1  MARHTGQGVDFQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF
51  GQIFRRTLIN GVVAVVLGFV VVKLGCGNDV YAGQPFVAVQD GAGIFAAADK
101 TFGNDFAXEG VSILRKRFSD GLFL*

```

**m096/g096** 96.0% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGI FEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN					
g096	MAGHTGQGVDFQQIEFAVGI FEEIDAHAAFRTDCLCAANRQFAHQAFFGFGQIFRRTLIN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDVYAGQPFVAVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD					
	:					
g096	GVVSVVLGFVVVKLGCGDDVYAGQPFVAVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD					
	70	80	90	100	110	120
m096.pep	GLFLX					
g096	GLFLX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 337>:

**a096.seq**

```

1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGCGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTG CGGTTGTACT
201 TGGCTTCGTA GTCATAAAAC TCGGTTGCGG GGATGATGTG TATGCCGGGC
251 AGCCCTTTGC CGTTCACAC AGGCAGGAA TATTCGCCGC CGCCGATAAA
301 CCCTTCGGCA ATGATTTCGC CCT.GAAAGT GTTCAATTTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 338; ORF 096.ng>:

**a096.pep**

```

1  MAGHTGQGVDFQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF

```

51 GQIFRRRLIN GVVAVVLGFV VIKLGRGDDV YAGQAFVQH RAGIFAAADK  
 101 PFGNDFAXES VSILRKRFSD GLFL\*

m096/a096 92.7% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGIFEEIDAHAAFR	TDC	LRAANRQFAHQAFFGFGQIFRRRLIN			
a096	MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFR	TDC	LRAANRQFAHQAFFGFGQIFRRRLIN			
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDVYAGQPFVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD					
a096	GVVAVVLGFVVVKLGCGNDVYAGQPFVQDGAGIFAAADKTFGNDFAXESVSILRKRFSD					
	70	80	90	100	110	120
m096.pep	GLFLX					
a096	GLFLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 339>:

g097.seq

```

1  ATGGATATTT CAAACAAAC ATTGCTGGAT AGGGTTTTTA ACCTGAAGGC
51  AAACGGTACG ACGGTACGTA CCGAGTTGAT GGCGGGTTTG ACGACCTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAAIC CCCTGATTTT GGGCGAGACC
151 GGAATGGATA TGGGGCGCGT ATTCGTCGCT ACCTGTATCG CATCCGCCAT
201 CGGCTGTTTT GTCATGGGTT TTATCGGCAA CTATCCGATT GCGCTTGCCC
251 CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
301 GGCGTGCCTT GGCAGGTGGC GTTGGGTGCG GTGTTCATTT CCGGTCTGAT
351 TTTCATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCCTATGGG TTTGAAAATG TCGATTGCCG CCGGTATCGG TTTGTTTTTG
451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTCGGC TTGGGCGATA TTCATCAGCC CAGCGCACTG TTGGCATTGT
551 TCGGTTTTGT CATGGTGGTC GTATTGGGGT ATTTCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATTCTGAC GATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAACGAGTTT CACGGCGTGG TCGGCGAAGT ACCGGGCATT GCGCCGACCT
701 TTATGCAGAT GGATTTTAAA GGTCTGTTTA CCGTCAGCAT GGTCAGCGTG
751 ATTTTCGTCT TCTTCTTGGT CGATTGTTC GACAGTACCG GAACGCTGGT
801 CGGCGTATCC CACCGTGCCG GACTGCTGGT GGACGGTAAG CTGCCCGGCC
851 TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTTCAACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TGGCGTGTCT GATGTTCTCC CCATTGGCGA AAAGTGTTC GGTATTTGCC
1051 ACCGCGCCCG CACTGCTTTA TGTGCGCACG CAGATGCTCC GCAGTGCGAG
1101 GGACATTGAT TGGGACGATA TGAAGAAGC CGCGCCCGCG TTCCTGACCA
1151 TTGTCTTCAT GCCGTTTACC TATTCGATTG CAGACGGCAT CGCCTTCGGC
1201 TTCATCAGCT ATGCCGTGGT CAAACTTTTG TGTGCGCGGA CTGGGGACGT
1251 GCCGCCTATG GTATGGGTTG TTGCCGTATT GTGGGCATTG AAATTCTGGT
1301 ATTTGGGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 340; ORF 097.ng>:

g097.pep

```

1  MDISKQTLID RVFNLIKANGT TVRTELMAGL TTFLTMCYIV IVNPLILGET
51  GMDMGAVFVA TCIASAIGCF VMGFIGNYPI ALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFVMV VLG YFRVQGA
201 IIITILTITV IASLMGLNEF HGVVGEVPGI APTFMQMDFK GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACL MFS PLAKSVPVFA
351 TAPALLYVGT QMLRSARDID WDMTEAAPA FLTIVFMPFT YSIADGIAFG

```

401 FISYAVVKLL CRRTGDPVPM VVVAVLWAL KFWYLG\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 341>:

m097.seq

```

1 ATGGACACTT CAAAACAAAC ACTGTGGAC GGGATTTTTA AGCTGAAGGC
51 AAACGGTACK ACGGTGCGTA CCGAGTTGAT GCGGGGTTTG ACAACTTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAACC CTCyGATTTT GGGCGAGACC
151 GGCATGGATA TGGGGGCGGT ATTCGTGCGT ACCTGTATCG CGTCTGCCAT
201 CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
251 CGGGGATGGG GCTGAATGCC TATTTACCTT TTGCCGTCGT TAAGGGTATG
301 GCGGTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTTCATCT CCGGTCTGAT
351 TTTTATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCCTATGGG TTGAAAATG TCGATTGCTG CCGGTATCGG TTGTTTTTG
451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTCGGT TTGGGCGATA TTCATCAGCC GTCCGCGTGT TTGGCATTGT
551 TCGGTTTTGC TATGGTGGTC GTATTGGGAC ATTCCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATCTTGAC CATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAATGAATTT CACGGCATCA TCGGCGAAGT ACCGAGCATT GCGCCGACTT
701 TTATGCAGAT GGATTTTGAA GGCCTGTTTA CCGTCAGCAT GGTCACTGTG
751 ATTTTCGTCT TCTTCTTGGT CGATCTATTT GACAGTACCG GAACGCTGGT
801 CCGCATATCC CACCGTGCCG GGCTGCTGGT GGACGGTAAG CTGCCCGGCC
851 TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTCCACCAC GCCTTATGTG GAAAGCGCGG CCGGCGTATC
951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTCG CGCTTTTGCC
1051 ACCGCGCCCG CCCTGCTTTA TGTCGGCAGC CAGATGTCTC GCAGTGCAGG
1101 GGATATTGAT TGGGACGATA TGACGGAAGC CGCACCTGCG TCCTTGACCA
1151 TTGTTTTCAT GCCGTTTACT TATTCGATTG CAGACGGCAT CGCTTTCGGC
1201 TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGCA CCAAAGACGT
1251 TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
1301 ATTTGGGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 342; ORF 097>:

m097.pep

```

1 MDTSKQTLDD GIFKLKANGT TVRTELMAGL TFLTMCIYIV IVPXILGET
51 GMDMGAVFVA TCIASAIGCF VMGFVGNYP IALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHOPSAL LALFGFAMVV VLGHFRVQGA
201 IITILTITV IASLMGLNEF HGIIEVPSI APTFMQMDFE GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGIS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDPVPM VWIVAVLWAL KFWYLG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 097 shows 96.3% identity over a 436 aa overlap with a predicted ORF (ORF 097.ng)

from *N. gonorrhoeae*:

m097/g097

	10	20	30	40	50	60
m097.pep	MDTSKQTLDDGIFKLKANGTTVRTELMAGLTTFLTMCIYIVIVNPXILGETGMDMGAVFVA					
	:					
g097	MDISKQTLDDRNVNLKANGTTVRTELMAGLTTFLTMCIYIVIVNPLILGETGMDMGAVFVA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m097.pep	TCIASAIGCFVMGFVGNYP IALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL					
	:					
g097	TCIASAIGCFVMGFIGNYP IALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL					
	70	80	90	100	110	120
	130	140	150	160	170	180

301

m097.pep	FSFFKVR	EMLVNAL	PMGLKMS	IAAGIGL	FLALISL	KGAGII	VANPATL	VGLGDI	HQPSAL
g097	FSFFKVR	EMLVNAL	PMGLKMS	IAAGIGL	FLALISL	KGAGII	VANPATL	VGLGDI	HQPSAL
		130	140	150	160	170	180		
m097.pep	LALFGF	AMVVVL	GHFRVQ	AIITILT	ITITVIA	SLMGLN	EFHGI	IGEVPS	IAPTFM
g097	LALFGF	AMVVVL	GHFRVQ	AIITILT	ITITVIA	SLMGLN	EFHGI	IGEVPS	IAPTFM
		190	200	210	220	230	240		
m097.pep	GLFTVSM	VSIVFV	FFLVLD	LFDS	TGTLVG	ISHFAG	LLVDG	KLPRLK	RALLAD
g097	GLFTVSM	VSIVFV	FFLVLD	LFDS	TGTLVG	SHRAGL	LLVDG	KLPRLK	RALLAD
		250	260	270	280	290	300		
m097.pep	LGTSS	TPYVES	AAGVS	AGGRT	GLTAVT	VGVLML	ACL	MFSP	LAKSV
g097	LGTSS	TPYVES	AAGVS	AGGRT	GLTAVT	VGVLML	ACL	MFSP	LAKSV
		310	320	330	340	350	360		
m097.pep	QMLRS	ARDIDW	DDMTEA	APFLT	IVFMP	FTYSI	ADGIA	FGFIS	YAVVK
g097	QMLRS	ARDIDW	DDMTEA	APFLT	IVFMP	FTYSI	ADGIA	FGFIS	YAVVK
		370	380	390	400	410	420		
m097.pep	VWIVAV	LWALKF	WYLGX						
g097	VWVAV	LWALKF	WYLGX						
		430							

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 343>

a097.seq

1	ATGGACACTT	CAAAACAAAC	ACTGTTGGAC	GGGATTTT	TAAGCTGAAGGC
51	AAACGGTACG	ACGGTGCGTA	CCGAGTTGAT	GGCGGGTTT	GAACACTTTT
101	TGACGATGTG	CTACATCGTT	ATCGTCAACC	CTCTGATTTT	GGGCGAGACC
151	GGCATGGATA	TGGGGGCGGT	ATTCGTCGCT	ACCTGTATCG	CGTCTGCCAT
201	CGGCTGTTTT	GTTATGGGTT	TTGTCGGCAA	CTATCCGATT	GCACTCGCAC
251	CGGGGATGGG	GCTGAATGCC	TATTTACCTT	TTGCCGTCGT	TAAGGGTATG
301	GGCGTGCCCT	GGCAGGTTGC	GTTGGGTGCG	GTGTTTCATCT	CCGGTCTGAT
351	TTTCATCCTG	TTCAGCTTTT	TTAAAGTCAG	GGAAATGCTG	GTCAACGCAC
401	TGCCTATGGG	TTTGAAAATG	TGCATTGCTG	CCGGTATCGG	TTTGTTTTTG
451	GCACTGATTT	CCCTGAAAGG	CGCAGGCATT	ATCGTTGCCA	ATCCGGCAAC
501	CTTGGTCCGC	TTGGGCGATA	TTCATCAGCC	GTCCCGCTTG	TTGGCACTGT
551	TCGGTTTTGC	CATGGTGGTC	GTATTGGGAC	ATTTCCGCGT	TCAAGGCGCA
601	ATCATCATCA	CCATTTTGAC	GATTACCGTC	ATTGCCAGCC	TGATGGGTTT
651	GAACGAATTT	CACGGCATCA	TCGGCGAAGT	GCCGAGCATT	GCGCCGACTT
701	TTATGCAGAT	GGATTTTAAA	GGGTGTTTTA	CCGTGAGCAT	GGTCAGCGTG
751	ATTTTCGTCT	TTTTCTAGT	CGATCTGTTT	GACAGTACCG	GAACACTGGT
801	CGGTGTATCG	CATCGTGCCG	GACTGCTGGT	GGACGGTAAG	CTGCCCGGCC
851	TGAAACGCGC	ACTGCTTGCA	GACTCTACCG	CTATTGTGGC	AGGTGCGGCT
901	TTGGGTACTT	CTTCAACCAC	GCCTTATGTG	GAAAGTGCGG	CGGGCGTATC
951	GGCAGGCGGG	CGGACAGGTC	TGACGGCGGT	TACCGTCCGC	GTATTGATGC
1001	TCCGCTGCCT	GATGTTTTCA	CCTTTGGCGA	AAAGTGTTCC	CGCTTTTGCC
1051	ACCGCGCCCG	CCCTGCTTTA	TCGGGCACG	CAGATGCTCC	GCAGTGCGAG
1101	GGACATCGAT	TGGGACGATA	TGACGGAAGC	CGCACCCGCA	TTCTTGACCA
1151	TTGTCTTCAT	GCCGTTTACC	TATTCGATTG	CAGACGGCAT	CGCTTTCCGC
1201	TTTCATCAGTT	ATGCCGTGGT	TAAACTTTTA	TGCCGCCGCA	CCAAAGACGT
1251	TCCGCCTATG	GTATGGATTG	TTGCCGTATT	GTGGGCACTG	AAATTCTGGT

1301 ATTTGGGCTG A

This corresponds to the amino acid sequence &lt;SEQ ID 344; ORF 097.a&gt;:

a097.pep

```

1  MDTSKQTLDD GIFKLGKANGT TVRTELMAGL TTFLTMCIYIV IVNPLILGET
51  GMDMGAVFVA TCIAAIGCF VMGFVGNYP IALAPGMGLNA YFTFAVVKGM
101 GVP::QVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
201 IITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFK GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDVP PM VWIVAVLWAL KFWYLG*

```

m097/a097 99.3% identity in 436 aa overlap

m097.pep	10	20	30	40	50	60
	MDTSKQTLDDGIFKLGKANGTTVRTELMAGLTTFLTMCIYIVIVNPNXILGETGMDMGAVFVA					
a097	10	20	30	40	50	60
	MDTSKQTLDDGIFKLGKANGTTVRTELMAGLTTFLTMCIYIVIVNPLILGETGMDMGAVFVA					
m097.pep	70	80	90	100	110	120
	TCIAAIGCFVMGFVGNYP IALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL					
a097	70	80	90	100	110	120
	TCIAAIGCFVMGFVGNYP IALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL					
m097.pep	130	140	150	160	170	180
	FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL					
a097	130	140	150	160	170	180
	FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL					
m097.pep	190	200	210	220	230	240
	LALFGFAMVVVLGHFRVQGAIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFE					
a097	190	200	210	220	230	240
	LALFGFAMVVVLGHFRVQGAIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFK					
m097.pep	250	260	270	280	290	300
	GLFTVSMVSVIFVFFLVDLFDSTGTLVGISHRAGLLVDGKLPRLKRALLADSTAIVAGAA					
a097	250	260	270	280	290	300
	GLFTVSMVSVIFVFFLVDLFDSTGTLVGVS HRAGLLVDGKLPRLKRALLADSTAIVAGAA					
m097.pep	310	320	330	340	350	360
	LGTSSTTPYVESAGVSAGGRTGLTAVTVGVLMACLMFSPAKSVPAFATAPALLYVGT					
a097	310	320	330	340	350	360
	LGTSSTTPYVESAGVSAGGRTGLTAVTVGVLMACLMFSPAKSVPAFATAPALLYVGT					
m097.pep	370	380	390	400	410	420
	QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVP PM					
a097	370	380	390	400	410	420
	QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVP PM					
m097.pep	430					
	VWIVAVLWALKFWYLGX					
a097	430					
	VWIVAVLWALKFWYLGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 345>:

```
g098.seq
1  ATGACCGCCG ACGGTCTCTT CGTCGCTTTC AACTTCAATA CGTTTGCCGT
51  TGTGCGAATA TTGATACCAG TACAGCAGGA TGCTGCCGAG GCTGGCGATC
101 AGTTTGTCGG CGATGTCGCG CGCTTCGCTG TCGGGATGGC TTTCCGCTTC
151 GGGATGAACG CAGCCGAGCA TGGACACGCC GGTACGCATC ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAAGTG TCGCACTTCT TCAAACCTCGC
351 ATTTTGTGC CAAATTAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 346; ORF 098.ng>:

```
g098.pep
1  MTADGLFVAF NFNTFAVVRI LIPVQQDAAQ AGDQFVGDDA RFAVGMAFAF
51  GMNAAEHGHA GTHHVHRMGM CROAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQQV CDFFKLAFLC QIRMS*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 347>:

```
m098.seq
1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
51  TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
101 AGTTTGTCGG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTC
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTT GCGCACTTCT TCAAACCTCGC
351 ATTTTGTGC CAAATCAGAA TGTCGTAA
```

This corresponds to the amino acid sequence <SEQ ID 348; ORF 098>:

```
m098.pep
1  MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDDA RFTFRMAFTF
51  RMNAAQHGHA GTHYVHRMGM CROAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQQV GDFFKLAFLC QIRMS*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 098 shows 89.6% identity over a 125 aa overlap with a predicted ORF (ORF 098.ng) from *N. gonorrhoeae*:

```
m098/g098

      10      20      30      40      50      60
m098.pep  MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVGDDVARFTFRMAFTFRMNAAQHGHA
          |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
g098      MTADGLFVAFNFNTFAVVRI LIPVQQDAAQAGDQFVGDDVARFAVGMAFAFGMNAAEHGHA
          10      20      30      40      50      60

      70      80      90     100     110     120
m098.pep  GTHYVHRMGMCROAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQQV GDFFKLAFLC
          |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
g098      GTHYVHRMGMCROAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQQV CDFFKLAFLC
          70      80      90     100     110     120

m098.pep  QIRMSX
          |||||
g098      QIRMSX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 349>:

```
a098.seq
1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
51  TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
```

304

```

101 AGTTTGTGGG CGATGTGCGG CGCTTCACTT TCCGGATGGC TTTCACGTTT
151 AGGATGAACG CAGCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTT:A TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCAGCTTCT TCAAACTCGC
351 ATTTTGTGC CAAATCAGAA TGTCTGTA

```

This corresponds to the amino acid sequence <SEQ ID 350; ORF 098.a>:

a098.pep

```

1 MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDDVA RFTFRMAFTF
51 RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHCFE LGFISGQLEF
101 VGQMAVNQQV GDFFKLAFLC QIRMS*

```

m098/a098 100.0% identity in 125 aa overlap

	10	20	30	40	50	60
m098.pep	MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVGDDVARFTFRMAFTFRMNAAQHGYA					
a098	MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVGDDVARFTFRMAFTFRMNAAQHGYA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m098.pep	GTHYVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC					
a098	GTHYVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC					
	70	80	90	100	110	120
m098.pep	QIRMSX					
a098	QIRMSX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 351>:

g099.seq

```

1 ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTGGA
51 GCTGACGGGC AAACGGCAGG CGGGCATTAC TGCCACAGAC ATCGTGTGG
101 CACTGACCGA ATTCTTGCCT AAAGAGCGCG TGGTCGGGGC GTTTGTGCA
151 TTTTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGAGT TCGGCGCGAC TGCCGCCATG TCGCCATCG
251 ACGCGCAAA CATTGATTAT TTGAACTGA CCGGACGTGA CGACGCGCAG
301 GTGAAATTGG TGGAACCTA CGCCAAAACC GCAGGCTTAT GGGCAGGTGG
351 CTTGAAAACC GCCGTTTATC CGCGCGTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTTGCCACC
451 GCGGATTGG CGGCGAAAG GCTGGCGAAG CTTACGAAG AGCCTTCAGA
501 CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCGTGTA
551 CCAATACTTC CAACCCGCGC AACGTTGTCG CCGCCGCACT GTTGGCACGC
601 AATGCCAACC GCCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGGAATCTA TTTGAAAGAA GCAGGCTTGT
701 TGCCCGAAAT GGAAAACTC GGCTTCGGTA TCGTCGCTT CGCATGTACC
751 ACCTGTAAAC GCATGAgcgG CGCGCTcgaC CCGAAAATCC AACAAGAAAT
801 CATCGACCGC GAtttgtacg cCACCgCGT ATTGTcAGGC AACCgCAACT
851 TCGACGCGCG TATCCATCCG TATGCGAAAC AGGCTTTCTT CGCTTCGCTT
901 CCTTTGGTCG TTGCCTACGC ATTGGCAGGT AGCATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGGCCG CGAATCCGC CTGAAAGATA
1001 TCTGGCCGAC AGACGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051 CCGCAACAAT TCCGCGACAT TTATATCCCG ATGTCCGACA CCGGCACAGC
1101 GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGACCGATG TCCACCTACA
1151 TCCGCGCTCC GCCCTATTGG GAAGGCGCAC TGGCAGGGGA ACGTACATTA
1201 AGAGGTATGC GTCCGCGCGC GATTTTGCCC GACAACATCA CCACCGACCA
1251 CATCTCgcca tCCAATGCGA TTTTGGCCGG cagTGCCgca ggtgaATATT

```



```

1301 TGGCGAAAAAT GGGTTTGCCT GAAGAagaCT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTtcgt
1451 tggcacgcgT tgaacCAGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTT ATCATCATTG CCGGTGCGGA
1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC GCGAAGGGC GTGCGGCTGG
1601 CGGGTGTGGA AGCCATCGCC GCCGAAGGTT TCGAGCGCAT CCACCGCACC
1651 AACCTCATCG GCATGGGCGT CTGCGCGCTG CAATTCAAAC CCGGCACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTGCGCG
1751 AACGCACACC GCGGTGCGGC CTGACCCTCG TGATTACCG TAAAAACGGA
1801 GAAACCGTCG AAGTTCGGT TACCTGCCGC CCCGATACCG CAGAAGAAGC
1851 ATTGGTATAT GAAGCCGGCG GCGTATTGCA ACGGTTTGCA CAGGACTTTT
1901 TGAAGGGGAA CCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 352; ORF 099.ng>:

g099.pep

```

1 MLGRASMMRL PDIVGVELTG KRQAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRTISNM TPEFGATAAM FAIDAQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWAGGLKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAAKGLAK PYEPPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLKRK PWVKSSFAPG SKVAGIYLKE AGLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQOEIHDR DLYATAVLSE NRNFDGRIHP YAKQAFILAS
301 PLVVAYALAG SIRFDIENDV LGVADGREIR LKDIWPTDEE IDAIVAIEYVK
351 PQQFRDIYIP MSDTGTAQKA PSPLYDWRPM STYIRPPYVW EGALAGERTL
401 RGMREPAIIP DNITTDHISP SNAILAGSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501 ETYMRKQPL I IAGADYQG GSSRDWAAKG VRLAGVEAIA AEGFERIHR
551 NLIGMGVLP L QFKPGTNRHT LQLDGTETYD VVGERTPRCG LTLVIHRKNG
601 ETVEVPVTCR PDTAEEALVY EAGGVLRQFA QDFLEGNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 353>:

m099.seq

```

1 ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
51 GCTGAACGGC AAACGGCAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
101 CACTGACCGA GTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTTCTCGAA
151 TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGAGT TCGGCGCGAC TCGCGCATG TTCGTATTG
251 ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
301 GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
351 CTTGAAAAAC GCCGTTTATC CTCGCGTTT GAAATTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCAAGTAACC CGCATGCCCG TTTTGCGACC
451 GCCGATTTGG CGGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCGGA
501 CGGCCAAATG CCCGACGGCT CGGTCATCAT CGCCGCGATT ACCAGTTGCA
551 CCAACACTTC CAACCCGCGC AACGTTGTTG CCGCCGCGCT CTTGGCACGC
601 AATGCCAACC GTCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCGGGCCTGT
701 TGCCCGAAAT GGAAAACTC GGCTTCGGTA TCGTCGCCTT CGCCTGCACC
751 ACCTGCAACG GCATGAGTGG CCGCTGGAT CCGAAAATCC AGAAAGAAAT
801 CATCGACCGC GATTGTACG CCACCGCGT ATTATCAGGC AACCGCAACT
851 TCGACGGCCG TATCCACCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901 CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGTATCCGTT TCGATATTGA
951 AAACGACGTA CTCGCGTTG CAGACGCAA GGAAATCCGC CTGAAAGACA
1001 TTTGGCCTGC CGATGAAGAA ATCGATGCCG TCGTTGCCGA ATATGTGAAA
1051 CCGCAGCAGT TCCGCGATGT GTATGTACCG ATGTTTCGACA CCGGCACAGC
1101 GCAAAAAGCA CCCAGTCCGC TGTACGATTG GCGTCCGATG TCCACCTACA
1151 TCCGCCGTCC GCCTTACTGG GAAGGCGCGC TGGCAGGGGA ACGCACATTA
1201 AGAGGTATGC GTCCGCTGGC GATTTTGGCC GACAACATCA CCACCGACCA
1251 CCTCTCGCCG TCCAATGCGA TTTTGGCCGT CAGTGCCGCA GCGGAGTATT
1301 TGGCGAAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCT ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAAAA ACGAAGACGG CAGCGTCCGC CAAGGCTCGT
1451 TCGCCGCGCT CGAACCCGAA GGCGAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGTGCGGA

```

```

1551 CTATGGTCAA GCTTCAAGCC GCGACTGGGC TGCAAAGGC GTACGCCTCG
1601 CCGGCGTAGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACCTTATCG GCATGGGCGT GTTGCCGCTG CAGTTCAAAC CCGACACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTGGTCGGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTCG TGATTCACCG TAAAAACGGC
1801 GAACCGTTG AAGTTCCCGT TACCTGCTGC CTCGATACTG CAGAAGAAGT
1851 ATTGGTATAT GAAGCCGGCG GCGTGTGCA ACGGTTTGCA CAGGATTTTT
1901 TGAAGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 354; ORF 099>:

```

m099.pep
1  MLGRASMMRL PDIVGVELNG KRQAGITATD IVLALTEFLR KERVVGAFVE
51  FFGEGARSLs IGRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAAKGLAK PYEPPSDGQM PDGSVIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLKPK PWKSSFAPG SKVAEILKE AGLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFFLASP
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPADEE IDAVVAEYVK
351 PQQFRDVYVP MFDTGTAQKA PSPLYDWRPM STYIRRPYVW EGALAGERTL
401 RGMRLAILP DNITTDHLSP SNAILAVSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVKNEDGSVR QGSFARVEPE GETMRMWEAI
501 ETYMNKQPL I IAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHR
551 NLIGMGVLP LFQKPDNRHT LQDGTETD VVGERTPRCD LTLVIHRKNG
601 ETVEVPVTC LDTAEVLVY EAGGVLRFA QDFLEGNA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 099 shows 96.2% identity over a 639 aa overlap with a predicted ORF (ORF 099.ng) from *N. gonorrhoeae*:

m099/g099

	10	20	30	40	50	60
m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFFVFFGEGARSLs					
g099	MLGRASMMRLPDIVGVELTGKRQAGITATDIVLALTEFLRKERVVGAFFVFFGEGARSLs					
	10	20	30	40	50	60
m099.pep	IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT					
g099	IGDRATISNMTPEFGATAAMFAIDAQTIDYLKLTGRDDAQVKLVETYAKTAGLWAGGLKT					
	70	80	90	100	110	120
m099.pep	IGDRATISNMTPEFGATAAMFAIDEQTIDYLKLTGRDDAQVKLVETYAKTAGLWADALKT					
g099	IGDRATISNMTPEFGATAAMFAIDAQTIDYLKLTGRDDAQVKLVETYAKTAGLWAGGLKT					
	70	80	90	100	110	120
m099.pep	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEPPSDGQMPDGSVIAAI					
g099	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEPPSDGQMPDGAIVIAAI					
	130	140	150	160	170	180
m099.pep	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEPPSDGQMPDGSVIAAI					
g099	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEPPSDGQMPDGAIVIAAI					
	130	140	150	160	170	180
m099.pep	TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAEILKEAGLLPEMEKL					
g099	TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAGIYLKEAGLLPEMEKL					
	190	200	210	220	230	240
m099.pep	TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAEILKEAGLLPEMEKL					
g099	TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAGIYLKEAGLLPEMEKL					
	190	200	210	220	230	240
m099.pep	GFGIVAFAC <del>TT</del> TCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAF <del>FLASP</del>					
g099	GFGIVAFAC <del>TT</del> TCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAF <del>FLASP</del>					
	250	260	270	280	290	300
m099.pep	GFGIVAFAC <del>TT</del> TCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAF <del>FLASP</del>					
g099	GFGIVAFAC <del>TT</del> TCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAF <del>FLASP</del>					
	250	260	270	280	290	300
	310	320	330	340	350	360

307

m099.pep	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIEWPADEEIDAVVAEYVKPQQFRDYYVP
g099	PLVVAYALAGSIRFDIENDVLGVADGREIRLKDIEWPTDEEIDAIVA EYVKPQQFRDIYIP
	310 320 330 340 350 360
m099.pep	370 380 390 400 410 420
	MFDTGTAQKAPSPLYDWRPMSTYIRPPYWEGALAGERTLRGMRPLAILPDNITTDHLS
g099	MSDTGTAQKAPSPLYDWRPMSTYIRPPYWEGALAGERTLRGMRFPAILPDNITTDHISP
	370 380 390 400 410 420
m099.pep	430 440 450 460 470 480
	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNEDGSVR
g099	SNAILAGSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNEDGSVR
	430 440 450 460 470 480
m099.pep	490 500 510 520 530 540
	QGSFARVEPEGETMRMWEAIEITYMNRKQPLIIAGADYQGSSRDWAAKGVRLACVEAIV
g099	QGSLARVEPEGQTMRMWEAIEITYMNRKQPLIIAGADYQGSSRDWAAKGVRLACVEAIA
	490 500 510 520 530 540
m099.pep	550 560 570 580 590 600
	AEGFERIHRTNLIGMVLPLQFKPDTNRHTLQLDGTETYDVVGERTPRCDLTLVIHRKNG
g099	AEGFERIHRTNLIGMVLPLQFKPGTNRHTLQLDGTETYDVVGERTPRCGLTLVIHRKNG
	550 560 570 580 590 600
m099.pep	610 620 630 640
	ETVEVPVTCCLDTAEVVLVYEAGGVLRFAQDFLEGNAAX
g099	ETVEVPVTCRPTDAEALVYEAGGVLRFAQDFLEGNAAX
	610 620 630 640

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 355>:

a099.seq

```

1  ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
51  GCTGAACGGC AACCGGAAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
101 CACTGACCGA GTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTGTGCGAA
151 TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
251 ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
301 GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
351 CTTGAAAACC GCCGTTTATC CGCGCGTTTT GAAATTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTTGCGACC
451 GCCGATTGGC CCGCAAAGG CTTGGCTAAA CCTTACGAAG AGCCTTCAGA
501 CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCCTGTA
551 CCAATACTTC CAATCCGCGC AACGTTGTCT CCGCCGCGCT GTTGGCACGC
601 AATGCCAACC GCCTCGGCTT GCAACGCAAA CCTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCAGATCTGC
701 TGCCCGAAAT GGA AAAACTC GGCTTCGGTA TCGTTGCCTT CGCATGTACC
751 ACCTGTAACG GCATGAGCGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
801 CATCGACCGC GATTGTGACG CCACCGCGT ATTGTGAGG AACC GCAACT
851 TTGACGGCCG TATCCATCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901 CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGCATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGGCAA AGAAATCCGC CTGAAAGACA
1001 TTTGGCTTAC CGATGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051 CCGCAGCAAT TTCGCGACGT TTATATCCCG ATGTTGACGA CCGGCACAGC
1101 GCAAAAAGCA CCAAGCCCGC GTACGACTG GCGTCCAATG TCTACCTATA
1151 TCCGCCGCCC ACCTTACTGG GAAGGCGCAC TGGCAGGGGA ACGCACATTA
1201 AGCGGTATGC GTCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA

```

```

1251 TCTCTCGCCA TCCAATGCGA TTTTGGCAAG CAGTGCCGCA GCGCAATATT
1301 TGGCAAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGTGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTTTCGC
1451 TGGCACGCGT TGAACCCGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCIC ATCATCATTG CCGGCGCGGA
1551 CTACGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601 CCGGCGTGGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACTTGATCG GTATGGGCGT GTTGGCGCTG CAGTTCAAAC CGGGTACCAA
1701 CCGCCACACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTGCGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTTG TGATTACCG TAAAAACGGC
1801 GAGACCGTCG AAGTCCCCAT TACCTGCCGC CTCGATACCG CAGAAGAAGT
1851 GTTGGTATAT GAAGCCGGTG GCGTATTGCA ACGGTTTGCA CAGGATTTTT
1901 TGAAGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 356; ORF 099.a>:

a099.pep

```

1 MLGRASMMRL PDIVGVELNG KRKAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDEQTIIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAGKGLAK PYEEPSDQGM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLQRK PWVKSSFAPG SKVAEYLKE ADLLPEMEKL GFGIVAFACF
251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFILASP
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPTDEE IDAIVAEYVK
351 PQQFRDVYIP MFDTGTAQKA PSPLYDWRPM STYIRPPYVW EGALAGERTL
401 SGMRLAILP DNITTDHLSP SNAILASSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501 ETYMNKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHRF
551 NLIGMGVLP QFKPGTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
601 ETVEVPITCR LDTAEVLVY EAGGVLRFA QDFLEGNA*

```

m099/a099 97.5% identity in 639 aa overlap

m099.pep	10	20	30	40	50	60
	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS					
a099	10	20	30	40	50	60
	MLGRASMMRLPDIVGVELNGKRKAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS					
m099.pep	70	80	90	100	110	120
	IGDRATISNMTPEFGATAAMFAIDEQTIIDY LKLTGRDDAQVKLVETYAKTAGLWADALKT					
a099	70	80	90	100	110	120
	IGDRATISNMTPEFGATAAMFAIDEQTIIDY LKLTGRDDAQVKLVETYAKTAGLWADALKT					
m099.pep	130	140	150	160	170	180
	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDQMPDGSVIAAAI					
a099	130	140	150	160	170	180
	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAGKGLAKPYEEPSDQMPDGAVIDAAI					
m099.pep	190	200	210	220	230	240
	TSCTNTSNPRNVVAAALLARNANRLGLKRKPWVKSSFAPGSKVAEYLKEAGLLPEMEKL					
a099	190	200	210	220	230	240
	TSCTNTSNPRNVVAAALLARNANRLGLQRPWVKSSFAPGSKVAEYLKEADLLPEMEKL					
m099.pep	250	260	270	280	290	300
	GFGIVAFACFCTCNGMSGALDPKIQKEIIDRDLATAVLSGNRNFDRGRIHPYAKQAFILASP					
a099	250	260	270	280	290	300
	GFGIVAFACFCTCNGMSGALDPKIQKEIIDRDLATAVLSGNRNFDRGRIHPYAKQAFILASP					
m099.pep	310	320	330	340	350	360
	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPADEEIDAVVAEYVKPQQFRDVYVP					

a099	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIEWPTDEEIDAIVAEYVKPQQFRDVYIY	310	320	330	340	350	360
m099.pep	MFDGTGAQKAPSPLYDWRPMSYIRPPYWEALAGERTLRGMRPLAILPDNITTHLSP	370	380	390	400	410	420
a099	MFDGTGAQKAPSPLYDWRPMSYIRPPYWEALAGERTLSGMRPLAILPDNITTHLSP	370	380	390	400	410	420
m099.pep	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNEDGSVR	430	440	450	460	470	480
a099	SNAILASSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNEDGSVR	430	440	450	460	470	480
m099.pep	QGSFARVEPEGETMRMWEAIEITYMNRKQPLII IAGADYGGQSSRDWAAKGVRLAGVEAIV	490	500	510	520	530	540
a099	QGSLARVEPEGQTMRMWEAIEITYMNRKQPLII IAGADYGGQSSRDWAAKGVRLAGVEAIV	490	500	510	520	530	540
m099.pep	AEGFERIHRTNLIGMGVLPQFKPDTNRHTLQLDGTETYPVVGERTPRCDLTLVIHRKNG	550	560	570	580	590	600
a099	AEGFERIHRTNLIGMGVLPQFKPGTNRHTLQLDGTETYPVVGERTPRCDLTLVIHRKNG	550	560	570	580	590	600
m099.pep	ETVEVPVTCCLDTAAEEVLVYEAGGVLQRF AQDFLEGNAAX	610	620	630	640		
a099	ETVEVPITCRLDTAEEVLVYEAGGVLQRF AQDFLEGNAAX	610	620	630	640		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 357>:

```

g102.seq
1      AtgtCCGCCA  AAactccgtc  gctcttcggc  ggcgcgatga  Ttatcgccgg
5      gaaggttata  ggcgcAGgtA  tgttccccaa  ccccacgcgc  aacttggggg
101    acgggttgtaat  aggtcgtcgt  attgtcgtgc  tgtacacctc  gtttccattc
151    tcctccggcgc  ccctcatgat  ttfgaaagtc  aacacccata  acCCcggaga
201    ggcaAGttt   gatccATGg  tCAAAgacct  gctcgGaCGc  ggcctgaaca
251    tcatacaacgg  catcgcgcgc  tcagtgtTcC  tatacggctc  ttatcagcgc
301    tacattttat  tcggcgcgta  cctGACCGCC  AAAGGCAtcg  GCagCGcAGT
351    AGGCCGCCAA  ATTTCgtct  CCGTCGGACA  actcgtcttc  tTGGCAtCC
401    TCGCCTTTTG  CGTATGGCCA  TCCGCAGCCT  TGGTCGACGC  CTTTACGGCC
451    GTCCTCATCG  GCGGcATGGT  ATTAACCTTT  ATTTGGGCAA  CCGGCGCCCT
501    GGTTGCGCAT  GCGAAACCGT  CCGTCTCTCT  CGACACCCAA  GCCCCGCTCG
551    GCACCGGCTA  TCGATTATC  GCGGCCACCG  CCTGCGCCGT  TGCTCTCGCT
601    TCCTTCGGCT  TCCACGGCAA  CGTTTCCAGC  CTGCTCAAT  ACTTTAAAGG
651    CGACGgcCc   aaagtGcGA  aATCcatctg  gGcaggtaca  ttgTTCGCT
701    tggtaattta  cgtccTCTgg  caaacgcgca  tCaaagtcac  CTGCGcgc
751    aacgagttcg  cCCCcgrtat  tgccgcgcgag  aggcaactCT  CCGTCCTgaa
801    tgaaaaccTG  tccaaatctg  cccaaacgcg  cgatatggaT  aAaatattgt
851    ccctatttcc  ctacatggca  atgcgcacct  cctttttagg  cgTAACctta
901    ggcctgtttg  acaacatctc  cgacattctc  aaatggaacg  acagatatgc
951    cggcgcgggc  cccaaaaacg  tcgcgctgaa  ctctcctggc  CCCctgattt
1001   cctggctgct  cctccccacc  ggccttctta  ccgccattgc  tgcgtccgcg
1051   ctggcgcgca  ccgctctggga  ccaagGcatc  atccccgcca  tgcgtcctca
1101   cgtttcccc  caaaaaattg  cGcGaggcaa  gacttaataA  gtttaCGCGC
1151   cgttggtgct  gatgttagtc  ttccttttct  gcatcgccaa  catcgcgcca
1201   CAGGTATTGA  GccAAatGa  ACtcgtCccc  GTATTTAAG  GATAA

```

This corresponds to the amino acid sequence <SEQ ID 358; ORF 102.ng>:

g102.pep

```

1  MSAKTPSLFG  GAMIIAGKVI  GGMFPNPTA  NLGDGLIGSL  IVLLYTWFPP
51  SSGALMILEV  NTHNPRGASF  DTMVKDLLGR  GWNIIINGIAV  ALVLYGSTYA
101 YILVGGDLTA  KGIGSAVGGK  ISLTVGQLVF  FGILAFVCWA  SARLVDRFTG
151 VLIIGMVLTF  IWATGGLVAD  AKPSVLFDTQ  APVGTGYWIY  AATALPVCLA
201 SFGFHGNVSS  LLKYFKGDAP  KVAKSIWAGT  LVALVIYVLW  QTAIQSNLPR
251 NEFAPVIAAE  RQLSVLNETL  SKFAQTGDMD  KILSLFPYMA  IATSFGLVTL
301 GLFDNIADIF  KWNDSMSGRG  TKTVALNFLP  PLISWLLLP  GFFTAIGASG
351 LAATVWDQGI  IPAMLLYVSP  QKIGAGKTYK  VYGGWLMLV  FLFGIANIAA
401 QVLSQMEVLP  VFKG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 359>:

```

m102.seq
1  ATGCCCAACA  AAACCCCTTC  ACTGTTCCGC  GGC CGCATGA  TTATCGCCGG
51  CACGGTCATC  GCGCGAGGCA  TGTCTCGCAA  CCCGACCGCC  ACATCCGGCG
101 TATGGTTTAC  CGGCTCGCTG  GCCGTGTTGC  TGTACACCTG  GTTTTCTATG
151 CTTTCCAGCG  GCCTGATGAT  TTTGGAAATC  AACACCCATT  ATCCGCACGG
201 CGCAAGTTTC  GACACGATGG  TCAAAGACCT  GCTCGGACGC  GGCTGGAACA
251 TCATCAACCG  CATCGCCGTC  GCCTTCGTTT  TATACCTGCT  TACTTACGCT
301 TATATCTTCG  TCGGCGGCGA  CCTGACCGCC  AAAGGCTTAG  GCAGCGCGGC
351 AGGCGGCGAC  GTTCACTCA  CCGTCGGACA  ACTCGTCTTC  TTCGGCATCC
401 TCGCCTTTTG  CGTATGGGCA  TCCGCACGCT  TGGTCGACCG  CTTACCCGGC
451 GTCTTTATCG  GCGGCATGGT  ATTGACCTTT  ATTTGGGCGG  CCGCGGGGCT
501 GATTGCCGAT  GCCAAGCCGT  CCGTCCTCTT  CGATACCCAA  GCCCCGCGCG
551 GCACAACTA  CTGGATTAC  GCGGCCACCG  CCCTGCCCCG  CTGCTCGCT
601 TCCTTCGGCT  TCCACGGCAA  CGTCTCCAGC  CTGCTCAAAT  ACTTTAAAGG
651 CGACGCGCCC  AAAGTGGCTA  AATCCATCTG  GACGGGCACA  CTGATTGCGC
701 TGGTAATTTA  CGTCCTCTGG  CAAACCGCCA  TCCAAGGCAA  CCTGCCGCGC
751 AACGAGTTCG  CCCCCGTCAT  CGCCGCCGAA  GGGCAAGTCT  CCGTCCTCAT
801 CGAAACCCTG  TCCAAATTCG  CCCAAACCGG  CAATATGGAC  AAAATATTGT
851 CCCTGTTTTC  CTATATGGCG  ATCGCCACCT  CGTTTTTAGG  CGTAACGCTC
901 GGACTCTTCG  ACTACATCGC  CGACATCTTC  AAATGGAACG  ACAGCATCTC
951 CGGCCGCACC  AAAACCGCG  CGCTGACCTT  CCTGCCGCC  CTGATTTCCT
1001 CGCTGCTCTT  CCCCACCGGC  TTCGTTACCG  CCATCGGCTA  CGTCGGCCTG
1051 GCGGCAACCG  TCTGGACAGG  CATCATCCCC  GCCATGCTGC  TCTACCGTTC
1101 GCGCAAAAAA  TTCGGCGCAG  GCAAAACCTA  TAAAGTTTAC  GCGCGCTTGT
1151 GGCTGATGGT  TTGGGTCTTC  CTTTTCGGCA  TCGTCAACAT  CGCCGCACAG
1201 GTATTGAGCC  AAATGGAAC  CGTCCCGGTA  TTAAAGGAT  AA

```

This corresponds to the amino acid sequence <SEQ ID 360; ORF 102>:

```

m102.pep..
1  MPNKTPSLFG  GAMIIAGTVI  GAGMLANPTA  TSGVWFTGSL  AVLLYTWFSM
51  LSSGLMILEV  NTHYPHGASF  DTMVKDLLGR  GWNIIINGIAV  AFVLYLLTYA
101 YIFVGGDLTA  KGLGSAAGGD  VSLTVGQLVF  FGILAFVCWA  SARLVDRFTG
151 VLIIGMVLTF  IWAAGGLIAD  AKPSVLFDTQ  APAGTNYWIY  AATALPVCLA
201 SFGFHGNVSS  LLKYFKGDAP  KVAKSIWGT  LIALVIYVLW  QTAIQGNLPR
251 NEFAPVIAAE  QVSVLIETL  SKFAQTGNMD  KILSLFSYMA  IATSFGLVTL
301 GLFDYIADIF  KWNDSISGR  KTAALTFLPP  LISCLLPFTG  EVTAIGYVGL
351 AATVWTGIIP  AMLLYRSRKK  FGAGKTYKVY  GGLWLMVWVF  LFGIVNIAAQ
401 VLSQMEVLPV  FKG*

```

m102/g102 86.0% identity in 415 aa overlap

	10	20	30	40	50	60
m102.pep	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
g102	MSAKTPSLFGGAMIIAGKVIAGGMFPNPTANLGDGLIGSLIVLLYTWFPPSSGALMILEV					
	10	20	30	40	50	60
m102.pep	NTHYPHGASFDTMVKDLLGRGWNIIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD					
g102	NTHNPRGASFDTMVKDLLGRGWNIIINGIAVALVLYGSTYAYILVGGDLTAKGIGSAVGGK					
	70	80	90	100	110	120
m102.pep	VSLTVGQLVFFGILAFVCWASARLVDRFTGVLIIGMVLTFIWAAGGLIADAKPSVLFDTQ					
g102	ISLTVGQLVFFGILAFVCWASARLVDRFTGVLIIGMVLTFIWAATGGLVADAKPSVLFDTQ					
	130	140	150	160	170	180
m102.pep	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWGTGLIALVIYVLW					
g102	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWGTGLIALVIYVLW					
	190	200	210	220	230	240
m102.pep	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWGTGLIALVIYVLW					
g102	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWGTGLIALVIYVLW					

a102.seq

1	ATGCCACCA	AAACCCCTTC	ACTGTTCGGC	GGCGGATGA	TTATCGCCGC
51	CACGNTC	GGCGCAGTA	TGCTCGCAA	CCCgACCGC	ACATCCGGCG
101	TATGGTTAC	CGGCTCGCT	GCCGTGTC	TGTACACCTG	GTTTTCATG
151	CTCTCCAGC	GCGTGATG	TTCGGAAGT	AACACCCACT	ACCCCCACGG
201	CGCGANCTT	GACACCATG	TTAAAGACCT	GCTCGGACGG	AGCTGGAACA
251	TCATCAACGG	CATCTCGGTC	GCCCTCGTTT	TATACCTGCT	TACTTAGCT
301	TATATCTTC	TCGGCGCGA	CCTGACGCC	AAAGCTTAG	CGAGCGCGGC
351	AGGGGCAAT	GTTTCACTA	CGCTCGGACA	ACTCGTCTTC	TTCCGCAATC
401	TCGCCTTTTG	CGTATGGGA	TCCGCACTGT	TGGTCGACCG	ATTCAACAGC
451	GTCCTCATCG	GCGCACTGGT	ATTAACCTTT	ATTTGGGCAA	CCGGCGGCCT
501	GATTGCGGAT	GGCAAACTGC	CGCTCTCTTT	CGACACCCAA	CGCCCTACCG
551	GCACCAACTA	CTGGATTAT	GTCCGCCACG	CCCTGCCCGT	CTGCCTTGCG
601	TCATTCTCGT	TCCACGGCAA	CGTCTCCAGC	CTGCTCAAAT	ACTTTAAAGG
651	CGACGCGCCC	AAAGTGGCTA	AATCCATCTG	CAGGGGCACA	CTGATTGCGC
701	TGGTAATTTA	CGCTCTTGG	CAAACCGCCA	TCCAANGCAA	CTCTGCCGCGC
751	AACGAGTTTC	CCCCCGTGAT	TGCCGCCGAA	GGGCAAGTCT	CCGTCTNTGAT
801	TGAAACCCGT	TCCAAATTCG	CCCAAAACCGG	CAATATGGAC	AAAATATTGT
851	CCCTGTTTTC	CTATATGGCG	ATCGCCACCT	CGTTTTTAGC	CGTAACGCTC
901	GGACTCTTCG	ACTACATCGC	GACATCTCTT	AAATGGAACG	ACAGCGGTGC
951	CGGCCGCACC	AAAACCGCGC	CGCTGACCTT	CCTGCCGCCT	NTAATTTCCCT
1001	GCTTGCTCTT	CCCAACCGCG	TTTGTATACC	CCATCGGNTA	CGTCCGCCTC
1051	CGCGCAACCG	TCTGGACAGG	CATCATCCCC	GCCATGCTGC	TNTACCGTTG
1101	GCGCAAAAAA	TTCGGCGCAG	GCAAAACCTA	TAAAGTTTAC	GCGCGCTTGT
1151	GGCTGATGGT	TTGGGTCTTC	CTTTTCGGCA	TCNTCAACAT	CGCCGCACAN
1201	GTATTGAGCC	AAATGGAAC	CGTCCCCGTA	TTTAAAGGAT	AA
1202					

This corresponds to the amino acid sequence <SEO ID 362; ORF 102.a>:

a102.pep

1	MPTKTPSLFG	GAMIIAGTXI	GAGMLANPTA	TSGVWFTGSL	AVLLYTWFESM
51	LSSGLMLELV	NTHYPHGAXF	DTMVKDLLGR	SWNIINGIAV	AFVLYLLTYA
101	YIFVGGDLTA	KGLGSLAAGN	VSLTVGVLDF	FGILAFVCWA	SARLVDRETS
151	VLIGGMVLTf	IWATGGLIAD	AKPLVLFDTQ	APTGTNYWYI	VATALPVCLA
201	SFGFHGNVSS	LLKYFKGDAP	KVAKSIWTGT	LIALVIYVLW	QTAIQXNLPR
251	NEFAPVIAAE	GQVSXVIETL	SKFAQTGNMD	KILSLFSYMA	IATSFLGVTL
301	GLFDYIADIF	KWNDSVSGRT	KTAALTFLPP	XISCLLFPTG	FVTAIGYVGL
351	AATVWTGIIP	AMLLYRSRKL	FGAGKTYKVY	GGLWLMVWVF	LFGIXNIAAX
401	VLSSOMELPV	FGK*			

m102 / a102 95.9% identity in 413 aa overlap

312

m102.pep	10	20	30	40	50	60
	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLTYWFSMLSSGLMILEV					
a102	10	20	30	40	50	60
	MPTKTPSLFGGAMIIAGTXIGAGMLANPTATSGVWFTGSLAVLLTYWFSMLSSGLMILEV					
m102.pep	70	80	90	100	110	120
	NTHYPHGASFDTMVKDLLGRGWNIIINGIAVAFVLYLLTYAIFVGGDLTAKGLGSAAGGD					
a102	70	80	90	100	110	120
	NTHYPHGAXFDTMVKDLLGRSWNIIINGIAVAFVLYLLTYAIFVGGDLTAKGLGSAAGGN					
m102.pep	130	140	150	160	170	180
	VSLTVGQLVFFGILAFVCWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
a102	130	140	150	160	170	180
	VSLTVGQLVFFGILAFVCWASARLVDRFTSVLIGGMVLTFIWAAGGLIADAKLPVLFDTQ					
m102.pep	190	200	210	220	230	240
	APAGTNYWIIYAATALPVCLASFGFHGNSVLLKYFKGDAPKVAKSIWTGTLIALVIYVLW					
a102	190	200	210	220	230	240
	APTGTNYWIIYVATALPVCLASFGFHGNSVLLKYFKGDAPKVAKSIWTGTLIALVIYVLW					
m102.pep	250	260	270	280	290	300
	QTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
a102	250	260	270	280	290	300
	QTAIQGNLPRNEFAPVIAAEGQVSVXIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
m102.pep	310	320	330	340	350	360
	GLFDYIADIFKWNDSISGRKTAAALTFLPPLISCLLFPTGFVTAIGYVGLAATVWTGIIP					
a102	310	320	330	340	350	360
	GLFDYIADIFKWNDSVSGRKTAAALTFLPPLISCLLFPTGFVTAIGYVGLAATVWTGIIP					
m102.pep	370	380	390	400	410	
	AMLLYRSRKKFGAGKTYKVYGGGLWLMVWVFLFGIVNIAAQVLSQMELVPFKGX					
a102	370	380	390	400	410	
	AMLLYRSRKKFGAGKTYKVYGGGLWLMVWVFLFGIXNIAAXVLSQMELVPFKGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 363>:

g105.seq

```

1   Atgtccgcag aaaCATACAc acAAAtcggc tGGgtaggct taggGcaaat
51  gGgtctgcct atgGTAACGC GGCTCTTGGA CGGCGGCATC GAAgTCGGCG
101 TATACAACCG CTCGCCCCGAC AAAACTGCCC CCATCTCgc CAAAGGAGCA
151 AAAGTTTACG GCagcACCGC CGAACTCGTC CGCGCCTGCC CCGTCATTTT
201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251 GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301 ACCGAAAACC TCGCGTCAA AGCACTTGTC GAAGCCGAG GCGGACAGTT
351 TGCCGAAGCA CCCGTTTCCG GATCGGTCGG ACCCGCCACC AACGGCACAC
401 TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
451 ATATTTTCCC TTGTCGGCAA AAAAACCTTC CATTTCCGGC ATGTCGGCAA
501 AGGCTCGGGC GCGAACTCG TCTTGAATC GCTCTTAGGC ATTTTCGGCG
551 AAGCGTACAG CGAAGCGATG CTGATGGCG GGCAGTTCGG CATCGATACC
601 GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGA CTGCCTATGTT
651 TCAAACAAAA AAATCACTAT GGGCAAACCG TGAGTTCCCC CCTGCCTTTG
701 CACTCAAACA CGCTTCCAAA GAcctTAACC TCGcgtcCA AGAGCTTGAA
751 CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801 CAAAGCAGTT GAAGCCGGCT ACGGCGAACA GGACGTTTC GCGGTTTACC
851 TGAAATTGGC AGAACA CTGA

```



This corresponds to the amino acid sequence <SEQ ID 364; ORF 105.ng>:

```
g105.pep
1  MSAETYTQIG WVGLGQMGLP MVTRLDDGGI EVGVYNRSPD KTAPISAKGA
51  KVYGSTAEVLV RACPVIPLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151 IFSLVGKKTFF HFGDVGKGSF AKLVNLSLLG IFGEAYSEAM LMARQFGIDT
201 DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 365>:

```
m105.seq
1  ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGaTAGGCT TAGGGCAAAT
51  GGGTCTGCTT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
101 TATACAACCG CTCGCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
151 AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251 GCGACGGATT GGCCGGCAAm ATCATCGTCA ACATGAGCAC CATCTCCCGG
301 ACCGAAaAGC TCGCCGTCAA AGCACTTGTC GAAGCGCAGm GaCAGTTTGC
351 CGAAGCACCC GTTTCGGGAT CGGTCGGGCC CGCCACCAAC GGCACGCTGC
401 TGATTCTGTT CGGCGGCAGC GAACCGtTTT AAACCCGCTG CAAAAAATAT
451 TTCCCTCGT CCGCAAAAAA ACCTTCCATT TCGGCGATGT CGGCAAAGGT
501 TCGGCGCGA AACTCGTCTT GAACTCGTC TTGGGCATT TCGGCGAaCG
551 TAcAGCGAAs GmTgCTGATG GCGCGGCAGT TCGGCATCGA TACCGACACC
601 ATCGTCGAAG CCATCGGGA CTGGCAATG GACTCGCCCA TGTTCCAAAC
651 CAAAAAATCC CTGTGGGCAA ACCGCGAATT CCGmCCGmC TTCGCCCTCA
701 AACACGCCTC CAAAGACCTC AACCTCGCCG TCAAAGAGCT TGAACAGGCA
751 GGCAACACCC TGCCCGCCGT CGAAACCGTT GGTGCCAGCT ACCGCAAAGC
801 AGTCGAAGCC GGCTACGGGA CACAGGACGT TTCCGCGCTT TACCTGAAAC
851 TGGCAGAACA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 366; ORF 105>:

```
m105.pep
1  MSANEYAQIG WIGLGQMGLP MVTRLDDGGI EVGVYNRSPD KTAPISAKGA
51  KVYGNTAEVLV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGX IIVNMSTISP
101 TEKLAVKALV EAQRQFAEAP VSGSVGPATN GTLLILFGGS EFPXTRCKKY
151 FPSSAKKPSI SamsakVRAR NSSXTRSWAF SANVQRXXLM ARQFGIDTDT
201 IVEAIGDSAM DSPMFQTKS LWANREFPXX FALKHASKDL NLAVKELEQA
251 GNTLPAVETV AASYRKAVEA GYGTQDVSGV YLKLAEH
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 105 shows 79.9% identity over a 289 aa overlap with a predicted ORF (ORF 105.ng) from *N. gonorrhoeae*:

```
m105/g105
10 20 30 40 50 60
g105.pep MSAETYTQIGWVGLGQMGLPMVTRLDDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAEVLV
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m105 MSANEYAQIGWIGLGQMGLPMVTRLDDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAEVLV
10 20 30 40 50 60

70 80 90 100 110 120
g105.pep RACPVIPLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m105 RDYPVIFLMVSDYAAVCDILNGVRDGLAGXIIVNMSTISPTTEKLAVKALVEAQR-QFAEA
70 80 90 100 110

130 140 150 160 170 180
g105.pep PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSAGKLVNLSLLG
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m105 PVSGSVGPATNGTLLILFGGSEFPXTRCKKYFPSSAKKP-SISAMSAKVRARNSSXTRSW
120 130 140 150 160 170
```

```

      190      200      210      220      230      240
g105.pep  IFGEAYSEANLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
           |:  ::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
m105      AFSANVQRXXLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPXXFALKHASK
           100      190      200      210      220      230

      250      260      270      280      289
g105.pep  DLNLAVKELEQAGNTLPAVETVAASRYKAVEAGYGEQDVSQVYLKLAEH
           |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
m105      DLNLAVKELEQAGNTLPAVETVAASRYKAVEAGYGTQDVSQVYLKLAEH
           240      250      260      270      280

```

a105.seq

1	ATGTCGCGAA	ACGAATACAC	ACAAATCGGC	TGGATAGGCT	TAGGGCAAT
51	GGGTCTGCCT	ATGGTAACGC	GGCTCTTGGA	CGGGCGCATC	GAAGTCGGCG
101	TATACAACCG	CTCGCCCGAC	AAAACCTGCC	CCATCTCCGC	CAAAGCGCTA
151	AAAGTTTACG	GCAACACCGC	CGAACTCGTC	CGCGACTATC	CCGTCAATTT
201	CTGTATGGTT	TCCGACTATG	CCGCCGTGTG	CGACATCTTG	AACCGAGTCC
251	GCGACGGATT	GGCCGCGCAA	ATCATCGTCA	ACATGAGCAC	CATCTCCCCG
301	ACCGAAAACC	TCGCCCTCAA	AGCATCTGTC	GAAGCCGCAG	GCGGCACGTT
351	TGCCGAAGCA	CCCGTTTCCG	GATCGGTGCG	GCCCGCCACC	AACGGCAGCT
401	TGCTGATTCT	GTTTCGCGCG	AGCGAAGCCG	TTTTAAACCC	GCTGCAAAAA
451	ATATTTTCCC	TCGTGCGCAA	AAAAACCTTC	CATTTCGGCG	ATGTCGCGCA
501	AGGTTCCGGC	GCGAAACTCG	TCTTGAACTC	GCTCTTGGGC	ATTTTCGGCG
551	AAGCGCTACG	CGAAGCGCAT	CTGATGGCGC	GGCAGTTCCG	CATCGATACC
601	GACACCATCG	TCGAAGCCAT	CGCGGGCTCG	GCAATTGCAT	CGCCCATGTT
651	CCAAACCAAA	AAATCCCTGT	GGGCAAAACG	CGAATCCCA	CCGCCCTTCG
701	CCCTCAAACA	CGCCTCCAAA	GACCTCAACC	TCGCCGTCAA	AGAGCTTGAA
751	CAGGCAGGCA	ACACCCTGCC	CGCGCTCGAA	ACCGTTGCTG	CCAGCTACCG
801	CAAGACGATC	GAAAGCCGGT	ACGGCGAACA	GGACGTTTCC	GGCGTTTACC
851	TGAAATTGGC	AGAACCATGA			

a105.pap

1	MSANEYTGIG	WIGLGQMLP	MVTRLDDGGI	EVGVYNRSPD	KTAPISAKGA
51	KYVGNTEALV	RDYPPVIFLMV	SDYAACVDCIL	NGVRDGLAGK	IIVNMSTPSK
101	TENLAVKALV	EAAAGQFAEA	SPVSGVGPAT	NGTLLILFGG	SEAVLNLTQPK
151	IFSLVGKKT	HFGDVGKGS	AKLVNLSLLG	IFGEAYSEAM	LMARQFGIDT
201	DTIVEAIGGS	AMDSPMFQTK	KSLWANREFP	PAFALKHAHK	DLNLAVKELE
251	QAGNTLPGVE	TVAASRYKAV	EAGYGEODVS	GVYLLKAEH*	

m105/a105 96.5% identity in 289 aa overlap

	10	20	30	40	50	60
m105.pep	MSANEYAQIGWIGLQMG	LPMTVTRL	LDGGIEVGVY	NRSPDKTAPISAKGAKVY	NGNTAE	LV
a105	MSANEYTIQIGWIGLQMG	LPMTVTRL	LDGGIEVGVY	NRSPDKTAPISAKGAKVY	NGNTAE	LV
	10	20	30	40	50	60
m105.pep	RDYPVIFL	MVSDYAAVCDIL	NGVRDGLAGKII	VNMSTISPTENLAVKALVEAAG-QFAEA		
a105	RDYPVIFL	MVSDYAAVCDIL	NGVRDGLAGKII	VNMSTISPTENLAVKALVEAAGGQFAEA		
	70	80	90	100	110	120
m105.pep	120	130	140	150	160	170
a105	120	130	140	150	160	170
	180	190	200	210	220	230

q105-1, seq

This corresponds to the amino acid sequence <SEQ ID 370; ORF 105-1.ng>:

q105-1.pap

1	MSAETYTQIG	WVGLQGMGLP	MVTRLDDGGI	EVGVYNRSPD	KTAPISAKGA
51	KVYVGSIAELV	RACPVIIFLMV	SDYAAVCDGL	NGVRDLFAGK	IVNNMSTIP
101	TENLAVKALV	AAAGGQFAEA	PSVGSVGPAT	NGTLLLFLGG	SEAVLNLPQK
151	IFSLVGKKTF	HFGDVGKGS	AKLVNLSLGL	IFGEAYSEAM	LMARQFGIDT
201	DTIVEAIGGS	AMDSPMFOKT	KSLWANREFF	PAFALKASHK	DLNLAVKLEL
251	OAGNTLPAVE	TVAASYRKFA	EAGYGEVNSP	GVYLKAFH*	

m105-1.seq

1	ATGTCGCGCAA	ACGAATACGC	ACAAATCGGC	TGGATAGGCT	TAGGGCAAAT
51	GGGTCTGCGCT	ATGCGTACCG	GGCTCTTGGA	CGCGCGGCAT	GAAGTCGGCG
101	TATACAAACG	CTGCGCCAGC	AAATCTGGCA	CCATCTCCGC	CAAGGCGCGCA
151	AAAGTTTACG	GCAACACCGC	CGAACTCGTC	CGCGACTATC	CCGTCATTTT
201	CTGTATGGTT	TCGCGATATG	CCGCGCGTGT	CGACATCCTG	AACGGAAGTC
251	GCAGCGAGAT	GGCGCGGCAA	ATCATCGTCA	ACATGAGCAC	CATCTCCCGC
301	ACCGAAAACC	TCGCGCTCAA	AGCACTTTGC	GAAGCCGCAG	GCGGACAGTT
351	TGCGCGAAGCA	CCCGTTTCCG	GATCGGTGCG	CGCCGCCACC	AACGCGACGC
401	TGCTGATTTG	TCTCGCGCGC	AGCGAAGCCG	TTTTCAACCC	GTGCGAAAAA
451	ATATTTTCCC	TCGTCGGCAA	AAAAACCTTC	CATTTTCGGC	ATGTCGGCAA
501	AGGTTTCGGC	GCGAANACTC	CTTTGAATCT	GCTCTTTGGC	ATTTTTCGGC
551	AAGCGTACAG	GCAGAACATG	TCTGTGGCGC	GGCAGTTGGG	CATCGATACC
601	GACACCATCG	TCGAAGCCAT	CGGgGACTCG	GCAATGGACT	CGCCCATGTT
651	CCAAACCAAA	AAATCCCTGT	GAGCAAAACG	CGAATTCCCG	CCGCGCTTGC
701	CCCTCAAAAC	GGCTCTCAAA	GAGCTCAAC	TCGCGGTGAA	AAGAGTTGAA
751	CAGGCAGGCA	ACACCTGCC	CGCCGTCGAA	ACCGTTGCTG	CCAGTACCG
801	CAAGACGAGT	GAGACCGGCT	ACGGCGAACA	GGACGTTTCC	GGCGTTTACC
851	TGAACCTGGC	GAAGCACTGA			

m105-1.pgp

1 MSANEYAQIG WIGLGQMGLP MVTRLDDGGI EVGVYNRSPD KTAPISAKGA  
51 KVGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP

316

```

101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFVG SEAVLNPLQK
151 IFSLVGKKTf HFGDVGKGSg AKLVNLSLLG IFGEAYSEXM LMARQFGIDT
201 DTIVEAIGDS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAELH*

```

m105-1/g105-1 96.9% identity in 289 aa overlap

```

          10      20      30      40      50      60
m105-1.pep MSANEYTAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNIAELV
          10      20      30      40      50      60
g105-1      MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV

          70      80      90     100     110     120
m105-1.pep RDYPVIFLMVSDYAACDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
          70      80      90     100     110     120
g105-1      RACPVIFLMVSDYAACDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA

          130     140     150     160     170     180
m105-1.pep PVSGSVGPATNGTLLILFVGSEAVLNPLQKIFSLVGKKTfHFGDVGKGSgAKLVNLSLLG
          130     140     150     160     170     180
g105-1      PVSGSVGPATNGTLLILFVGSEAVLNPLQKIFSLVGKKTfHFGDVGKGSgAKLVNLSLLG

          190     200     210     220     230     240
m105-1.pep IFGEAYSEXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK
          190     200     210     220     230     240
g105-1      IFGEAYSEAMLARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK

          250     260     270     280     290
m105-1.pep DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSgVYLKLAELH
          250     260     270     280     290
g105-1      DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSgVYLKLAELH

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 373>:

a105-1.seq

```

1   ATGTCCGCAA ACGAATACAC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
51  GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGCGCGCATC GAAGTCGGCG
101 TATACAACCG CTCGCCGAC AAAACTGCC CCATCTCCG CAAAGGCGCA
151 AAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
251 GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCG
301 ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCG GCGGACAGTT
351 TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC
401 TGCTGATTCT GTTCGGCGGC AGCGAAGCG TTTTAAACCC GCTGCAAAA
451 ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
501 AGGTTTCGGC GCGAAACTCG TCTTGAATC GCTCTTGGG ATTTTCGGCG
551 AAGCGTACAG CGAAGCGATG CTGATGGCG GGCAGTTCG CATCGATACC
601 GACACCATCG TCGAAGCCAT CGCGGGCTCG GCAATGGACT CGCCCATGTT
651 CCAAAACAAA AAATCCCTGT GGGCAAAACG CGAATTCCCA CCCGCTTCC
701 CCCTCAAAAC CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
751 CAGGCAGGCA ACACCTGCC CGCCGTGCAA ACCGTTGCTG CCAGCTACCG
801 CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTCC GCGGTTTACC
851 TGAAATTGGC AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 374; ORF 105-1.a>:

a105-1.pep

```

1   MSANEYTAQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPlSAGKA
51  KVGYNIAELV RDYPVIFLMV SDYAACDIL NGVRDGLAGK IIVNMSTISP
101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFVG SEAVLNPLQK
151 IFSLVGKKTf HFGDVGKGSg AKLVNLSLLG IFGEAYSEAM LMARQFGIDT
201 DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAELH*

```

a105-1/m105-1 99.0% identity in 289 aa overlap

```

          10      20      30      40      50      60
a105-1.pep MSANEYTAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNIAELV
          10      20      30      40      50      60

```

317

m105-1	MSANEYAQIGWIGLGMGLPMVTRLDDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
	10 20 30 40 50 60
a105-1.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
	70 80 90 100 110 120
m105-1	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
	70 80 90 100 110 120
a105-1.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKSGAKLVLSLLG
	130 140 150 160 170 180
m105-1	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKSGAKLVLSLLG
	130 140 150 160 170 180
a105-1.pep	IFGEAYSEAMLMARQFGIDTDITVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
	190 200 210 220 230 240
m105-1	IFGEAYSEAMLMARQFGIDTDITVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
	190 200 210 220 230 240
a105-1.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAHX
	250 260 270 280 290
m105-1	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAHX
	250 260 270 280 290

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 375>:

```

g107.seq
1  ATGGTATTAA CCTTTATTG GGCACCGGC GGCCTGGTTG CCGATGCCAA
51  ACCGTCCGTC CTCTTCGACA CCCAAGCCCC CGTCGGCACC GGCTACTGGA
101 TTTACGCCGC CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG cgcCcaaagt
201 GgCGAaATCc atctggGcag gtacattggt TGCCttggtta atttacgtcc
251 TCTggcaaac cgccatCcaa agcaaccTGC cgcgcaacga gttcgcCCCC
301 gtgattgccc cggagaggca actCTCCGTC CTgaatgaaa cccTGtccaa
351 attcgcccaa accggcgata tggataAaat attgtcccta ttccctaca
401 tggcaatcgc cacctccttt ttaggcgTAA Ccttaggcct gtttgacaac
451 atcgccggac atcttcaaat ggaacgacag tatgtccggg cgcaccacaaa
501 accgtcgcgc tga

```

This corresponds to the amino acid sequence <SEQ ID 376; ORF 107.ng>:

```

g107.pep
1  MVLTFIWATG GLVADAKPSV LFDTQAPVGT GYWIYAATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWAGTLVALV IYVLWQTAIQ SNLPRNEFAP
101 VIAAERQLSV LNETLSKFAQ TGDMDKILSL FPYMAIATSF LGVTLGLFDN
151 IAGHLOMERQ YVRAAPKFSR *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 377>:

```

m107.seq
1  ATGGTATTGA CCTTTATTG GCGCGCCGCG GGGCTGATTG CCGATGCCAA
51  GCGTCCGTC CTCTTCGATA CCCAAGCCCC CGCCGGCACA AACTACTGGA
101 TTTACGCCGs CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTCATCGCCG CCGAAGGGCA AGTCTCCGTC CTCATCGAAA CCCTGTCCAA
351 ATTCGCCCAA ACCGGCAATA TGGACAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCCATC TTCAAATGGA ACGACAGCAT CTCCGGgCCG CACCAAAACC
501 GCCGCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 378; ORF 107>:

```

m107.pep..
1  MVLTFIWAAG GLIADAKPSV LFDTQAPAGT NYWIYAXTAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP

```

101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTGLGLFDY  
151 IAHLMERQH LRAAPKPPR\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 107 shows 89.4% identity over a 170 aa overlap with a predicted ORF (ORF 107.ng)

from *N. gonorrhoeae*:

m107/g107

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	:     :     :     :     :     :					
g107	MVLTFIWATGGLVADAKPSVLFDTQAPVGTGYWIYAATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
	:     :     :     :     :     :					
g107	KGDAPKVAKSIWAGTLVALVIYVLWQTAIQSNLPRNEFAPVIAAERQSLVNETLSKFAQ					
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSFLGVTGLGLFDYIA-HLQMERQHLRAAPKPPR					
	:     :     :     :     :     :					
g107	TGDMDKILSLFPYMAIATSFLGVTGLGLFDNIAGHLQMERQYVRAAPKPPSR					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 379>:

a107.seq

```

1  ATGGTATTAA CCTTTATTG GGCACCGGC GGCCTGATTG CCGATGCCAA
51  ACTGCCCCTC CTCTTCGACA CCAAGCCCC TACCGGCACC AACTACTGGA
101 TTTATGTCGC CACCGCCCTG CCGCTCTGCC TTGCGTCATT CGGTTCCAC
151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGCCTGGTA ATTTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTGATTGCCG CCGAAGGGCA AGTCTCCGTC CTGATTGAAA CCCTGTCCAA
351 ATTGCGCCAA ACCGCAATA TGGACAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTGCGACTAC
451 ATCGCCGACA TCTTCAAATG GAACGACAGC GTGTCCGGCC GCACCAAAAC
501 CCGCCGCGTG ACCTTCCTGC CGCCTCTAAT TTCCTGCCTG CTCTTCCCCA
551 CCGGCTTTGT TACCGCCATC GGCTACGTCG GCCTGGCGGC AACCGTCTGG
601 ACAGGCATCA TCCCGGCCAT GCTGCTCTAC CGTTCGCGCA AAAAATTCCG
651 CGCAGGCAAA ACCTATAAAG TTTACGGCGG CTTGTGGCTG ATGGTTTGGG
701 TCTTCCTTTT CGGCATCGTC AACATCGCCG CACAGGTATT GAGCCAAATG
751 GAACTCGTCC CCGTATTAA AGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 380; ORF 107.a>:

a107.pep

```

1  MVLTFIWATG GLIADAKLPV LFDTQAPTGT NYWIYVATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP
101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTGLGLFDY
151 IADIFKWNDS VSGRTKTAAL TFLPPLISCL LFPTGFVTAI GYVGLAATVW
201 TGIIPAMLLY RSRKKFGAGK TYKVYGGWLW MVWVFLFGIV NIAAQVLSQM
251 ELVPVFKG*

```

m107/a107 94.8% identity in 154 aa overlap

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	:     :     :     :     :     :					
a107	MVLTFIWATGGLIADAKLPVLFDTQAPTGTNYWIYVATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60

319

```

          70      80      90      100      110      120
m107.pep  KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a107      KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ
          70      80      90      100      110      120

          130     140     150     160     170
m107.pep  TGNMDKILSLFSYMAIATSLGLVTLGLFDYIAHLQMERQHLLRAAPKPPRX
          |||||||||||||||||||||||||||||
a107      TGNMDKILSLFSYMAIATSLGLVTLGLFDYIADIFKWNDSVSGRTKTAALTFLPPLISCL
          130     140     150     160     170     180

a107      LFPTGFVTAIGYVGLAATVWTGIIPAMLLYRSRKKFGAGKTYKVYGGGLWLMVWVFLFGIV
          190     200     210     220     230     240

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 381>:

```

g108.seq
1  ATGttgccgg gCTTCAACCG GATATTCAaa cggTTTGCTC CAACACTCGG
51  AACggCGCAT AAAACGCCgc ccTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATCCTA TCGCCATAAA AGGCGGGGT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGCG GCAATCTTAC TCGGCGGGTG
201 CGCCGCCGGC GGCAACACAT TCGGCAGCTT AGACGGCGGC ACGGGTATGG
251 GTGGCAGCAT CGTCAAAATG ACGGTAGAAA gccAATGCCG TCGGGAATTG
301 GACAGGCGCA GCGAATGGCG TTTGACCGCG CTGGCGATGA GTGCCGAAAA
351 ACAGGCGGAA TGGGAAAACA AGATTGCGG CTGCGCTACC GAAGAAGCAC
401 CTAACCACT GACCGGCAAC GATGTGATGC AGATGCTGAa ccagtccacG
451 CGCaatcagg cacTtgccgc CctgaccgTC AAAacgggTT CcgctgcTT
501 CAaacgcctg tACCGCTAa

```

This corresponds to the amino acid sequence <SEQ ID 382; ORF 108.ng>:

```

g108.pep
1  MLPGFNRIFK RFAPTLGTAH KTPPFALSRT GRLIRSYRHK RRGFNRKGIE
51  MNKTLNILPA AILLGGCAAG GNTFGSLDGG TGMGGSIVKM TVESQCRAL
101 DRRSEWRLTA LAMSAEKQAE WENKICGCAT EEAPNQLTGN DVMQMLNQST
151 RNQALAAALTV KTVSACFKRL YR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 383>:

```

m108.seq
1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTT CAACACTCGG
51  AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATTCTA TCGCCATAAA AGGCGGGGT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGT GCAATCTTAC TCGGCGGCTG
201 CGCCGCCGGA GGCGGTAACA CATTCCGCAG CTTAGACGGT GGCACAGGCA
251 TGGCGGCAG CATCGTCAA ATGGCGGTTG GGAGCCAATG CCGTGCAGAA
301 TTGGACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351 AAAACAGGCG GAGTGGGAAA ACAAGATTG CGCTTGCCTC GCCCAAGAAG
401 CACCCGAACG GATGACCGGC AACGATGTGA TGCAGATGCT GGCTCCGTCC
451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
501 CTTCAAACAC CTGTACCGCT AA

```

This corresponds to the amino acid sequence <SEQ ID 384; ORF 108>:

```

m108.pep
1  MLPGFNRIFK RFVPTLGTAH KTPPFALSRT GRLIRFYRHK RRGFNRKGIE
51  MNKTLNILPV AILLGGCAAG GGNTFGSLDG GTMGGSIVK MAVGSQRAE
101 LDKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPERMTG NDVMQMLAPS
151 TRNQALAAALT AKTVSACFKH LYS*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 108 shows 89.6% identity over a 173 aa overlap with a predicted ORF (ORF 108.ng) from *N. gonorrhoeae*:

m108/g108

[illegible]

a108.seq

```

1   ATGTTGCCGG  GCTTCAACCG  GATATTCAA  CGGTTTGTT  CAACACTCGG
51  AACGGCGCAT  AAAACGCGCG  CCTTCGCGT  ATCCCGAAG  GGGCGGCTAA
101 TCAGATTCTA  TCGCCATAAA  AGGCGGGGT  TCAACCGAA  AGGAATTGAG
151 ATGAATAAAA  CCTTGTCAT  TTTGCGGTT  GCAATCTTA  TCGGCGGCTG
201 CGCGCGCGGG  GCGGTGTACA  CATTGCGCAG  CTTAGACGG  GGCACAGGTA
251 TGGGCGGCAG  CATCGTCAA  ATGCGGGTAG  AAAGCCAAT  CCGTGCGGAA
301 TTGAACAAAC  GCAGCGAAT  GCGTTTGACC  GCGTGGGCA  TGAGTGCCGA
351 AAAACAGGCG  GAATGGGAA  ACAAGATTG  CGTTGCGTC  GCCCAAGAAG
401 CACCCAACCA  CTGACCGGC  AACGATGTGA  TGCAGATGT  GGATCCGTCC
451 ACGCGCAATC  AGGCACTTGC  GCCTTGACC  GCCAAAACGG  TTTCCGCCTG
501 CTTCAAACAC  CTGTACCGCT  AA

```

a108.pcp

```

1  MLPGFNRIKF RFVPTLTGAH KTPPFALSRT GRLIRFYRHK RRGFNKRGIE
51 MNKTLSLILPV AILGGCAAG GENTFGSLDG GTMGGSIVK MAVESQCRAE
101 LNKRSEWRLT ALAMSAEKQA EWENKICAV AQEAPNQLTG NDVMQMLDPS
151 LRNQALAAVT AKTVSAEKFH LYR*

```

	10	20	30	40	50	60
m108.pep	MLPGFNRI	FKRFVPT	LGTAHKTP	PFALSRT	GRLIRFYR	HKRRGFNR
a108	MLPGFNRI	FKRFVPT	LGTAHKTP	PFALSRT	GRLIRFYR	HKRRGFNR
	10	20	30	40	50	60
	70	80	90	100	110	120
m108.pep	AILLGCCA	AGGGNTF	GSLDGGT	GMGGSIV	KMAVGSQ	CRAELDKR
a108	AILLGCCA	AGGGNTF	GSLDGGT	GMGGSIV	KMAVESQ	CRAELNKR
	70	80	90	100	110	120
	130	140	150	160	170	
m108.pep	EWENKIC	ACVAQE	APERMTG	NDVQMQL	APSTRNQ	ALAAALT
a108	EWENKIC	ACVAQE	APNQLT	GNDVQMQL	DPSTRNQ	ALAAALT
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 387>:



g109.pep

1	MYRRVVGLS	DGLGLAAGI	DRRRMLTAFG	SGHGNDARQ	NHPIRRHRGV
51	LFRLVNPVFG	WALTMLLDNL	GLIGCKERSA	QLGFVGRVLI	<u>PAVGFLILCV</u>
101	AMGAVGMLPG	IPPFLEQFS	LG		

m109.seq

```
1 ATGTATTATC GCCGGTGTAT GGGGCTATCC GATGGACTTG GCGATTTGGC
51 AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCCGCTTTG
101 GAAGCGGGCA TGGAAATGAC TCGCAAAGGC AAAACCAACC AATCCGCGCG
151 CATCGTGGTG TPTCTTTCG CTTGTCAAT CCGGTTTTCG GCTGGGCGTT
201 GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAAA GAGCGCAGTG
251 CGCAATTAGG TTTCGCCGGA CGCGTGTGTA TACCCGCAGT AGGTTTCTTG
301 ATCTTGTAGT TGGCGATGGG TCGCGTCGGG ATGCTGCCCG GATCCCGGCC
351 GTTTTGTGAA CACTTCAAA CTTTGGGCTA G
```

m109.pep  
1 MYYRRVMGLS DGLGDLAAGI ERSLGRRRIL TAFGSGHGND AQRQNHPIRR  
51 HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFAG RVLIPAVGFL  
101 ILCVAMGAVG MLPGIPPFL HFKSLG\*

Homology with a predicted ORF from *N. gonorrhoeae*

m109/q109

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 391>:

a109.seq

1	ATGTATTATC	GCCGGGTGT	GGGGCTATCC	GATGGACTTG	GCGATTTGGC
51	AGCCGGTATT	GAGCGTAGCC	TTGGTCGTAG	GCGTATACIT	ACCGCTTTTG
101	GAAAGCGGGCA	TGGAAATGAC	GCGCAAAGGC	AAAACCAACC	AATCCGCCGC
151	CACCGTGGTG	TTCTCTCCG	CTTGGTCAAT	CCGGTTTTTC	GCTGGGGCTT

322

```

201 GACGATGCTG TTGGATAAATT TGGGCTTAAT CGGCTGCAAA GAGCGCAGCG
251 CGCAATTAGG TTTCACCGGA CGCGTATTGA TACCCGTAGT AGGTTTCTTG
301 ATCTTGTTGTG TGGCGATGGG TGGCGTCGGG ATGCTGCCCC GTATCCCGCC
351 GTTTTGGAG CACTTCAAAT CTTTGGGCTA G

```

This corresponds to the amino acid sequence <SEQ ID 392; ORF 109>:

a109.pep

```

1 MYYRRVVGLS DGLGDLAAGI ERSLGRRRIL TAFSGSHGND AQRQNHPIRR
51 HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFTG RVLIPVVGFL
101 ILCVAMGAVG MLPGIPPFLE HFKSLG*

```

m109/a109 97.6% identity in 126 aa overlap

m109.pep	10	20	30	40	50	60
	MYYRRVVGLS DGLGDLAAGIERSLGRRRIL TAFSGSHGND AQRQNHPIRRHRGVLFRLVN					
a109	MYYRRVVGLS DGLGDLAAGIERSLGRRRIL TAFSGSHGND AQRQNHPIRRHRGVLFRLVN					
	10	20	30	40	50	60
m109.pep	70	80	90	100	110	120
	PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE					
a109	PVFGWALTMLLDNLGLIGCKERSAQLGFTGRVLIPVVGFLILCVAMGAVGMLPGIPPFLE					
	70	80	90	100	110	120
m109.pep	HFKSLGX					
a109	HFKSLGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 393>:

g111.seq

```

1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
51 CCTGGGTTTC ATCTTCTCTGA ACGCCTGTTC GGAacaaacC GCGCAaaccg
101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
201 TGATGATGCG CTAAAGAAG TCAACCGGCA GATGtccaCC TACCAGACCG
251 ATTCGAAAT CAGCCGGTt atacagacan atgctggaga gctcttcgcg
301 tntcatgcag nttctataac tgattccgcc gaagactgtc tgcctaatac
351 gcctatctca tcggcgctct ga

```

This corresponds to the amino acid sequence <SEQ ID 394; ORF 111.ng>:

g111.pep

```

1 MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLOGET MGTTYTVKYL
51 SNNRDKLPSP AKIQKRIDDA LKEVNRQMST YQTDSEISRF IQTAGELFAH
101 ASITDSAEDC LPNTPISSAL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 395>:

m111.seq

```

1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGTCT TGATATTTGC
51 CCTGGGTTTC ATCTTCTCTGA ACGCCTGTTC GGAACAAACC GCGCAAAACCG
101 TTACCCTGCA AGGCGAAACG ATGGGCACGA CCTATATCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AWAACACGCAT
201 CGATGACGCG CTAAAGAAk TCAACCGGyA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG
351 CCTGACACAC GGC GCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCCG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
501 AGGCAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG CCTATTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGCGGCGAA

```

```
m111.pep
  1  MPSETRLPNF  IRVLIFALGF  IFLNACSEQT  AQTVTLQGET  MGTTYXVKYL
51  SNNRDKLPSP  AEIXKRIDDA  LKEXNRXMST  YQPDSEISRF  NQHTAGKPLR
101 ISSDFAHVTA  EAVRLNRLTH  GALDVTVGPL  VNLWGFGPDK  VSTRPSPEO
151 IKQAASYTGI  DKILKQGGK  YASLSKTHPK  AYLDLSSIAK  GFGVDKVAGE
201 LEKYGIQNYL  VEIGGELHGK  GKNARGEPPW  IGIEQPNIVQ  GGNQTQIVPL
251 NNRSLATSGD  YRIFHVDKNG  KRLSHIINPN  NKRPIISHNLA  SISVVADSAM
301 TADGLSTGLF  VLGETEALKL  AEREKLAFLV  IVRDKGGYRT  AMSSEFEKLL
351 R*
```

```

m111.pep/g111.pep
      10      20      30      40      50      60
m111.pep  MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLOGETMGTTYXVKYLSNNRDKLPSP
          ||||| |||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g111      MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLOGETMGTTYTVKYLSNNRDKLPSP
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      10      20      30      40      50      60
      70      80      90     100     110     120
m111.pep  AEIXKRIDDALKEYNRXMSTYQPDSEISRNFNQHTAGKPLRISSDFAHVHTAEAVRLNRLTH
          ||: ||||| ||||| ||||| ||||| ||||| ||||| |||||
g111      AKIQKRIDDALKEVNRQMSTYQTDSEISRFIQT XAGELFAXHAXSITDSAEDCLPNTPI S
          70      80      90     100     110     120
      130     140     150     160     170     180
m111.pep  GALDVTVGPLVNLWGFPGDPKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
g111      SALX

```

```

a111.seq
1   ATGCCGTCGTG AAACACGCCT GCCGAACCTT ATCCGCACCT TGATATTTGC
51  CCTGAGTTTTC ATCTTCTCTGA ACGCCTGTTC GGAACAAACC GCGCAAAACCG
101 TTACCTCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151 TCAAAATACTC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
201 CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
251 ACTCCGAACT CAGCCGGTTC AACCAACACA CAGCCGGGCA GCCCTCCCG
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCGCTCC ACCTGAAACCG
351 CTTGACACAC GCGCGCTGG ACGTAACCGT CGTCCCTTG CTCAACCTTT
401 GGGGATTCGG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAAG CAGCATCTTA TAGGGGCATA GACAAAATCA TTTTGAACA
501 AGGCAAAGAT TACGTTCTCT TGAGCAAAAC CCACCCCAAG GCCTATTTTG
551 ATTTATCTTC GATTGCCAAC GGTTCGGCG TTGATAAAGT TCGGGCGGAA
601 CTGAAAAAAT ACGGCATTCA AAATATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
701 AACAGCCCCA CATCTCCCAA TTCCGGCAAT CGCAGATTAT TCTCCCGCTG
751 AACACCGGTT CGCTTGCCAC TTCCGGCAT TACCGTATTT CGCAGCTCGA

```

```

801 TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
901 ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAAGCTGC TGTTCCTG ATTGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTGA AAAAGTCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 398; ORF 111.a>:

a111.pep

```

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLOGET MCTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMS TQPDSEISRF NCHTAGKPLR
101 ISSDFAHVTA EAVHLNRLTH GALDVTGVL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKIIKQGGK YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHKG GKNARGEPRW IGIEQPNIVQ GGNTQIIVPL
251 NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

m111/a111 97.7% identity in 351 aa overlap

m111.pep	10	20	30	40	50	60
	MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLOGETMCTTYXVKYLSNNRDKLPSP					
a111	MPSETRLPNFIRVLIFALSFIFLNACSEQTAQTVTLOGETMCTTYTVKYLSNNRDKLPSP					
	10	20	30	40	50	60
m111.pep	70	80	90	100	110	120
	AEIXKRIDDALKEVNRQMSYQPDSEISRFNCHTAGKPLRISDFAHVTA EAVRLNRLTH					
a111	AEIQKRIDDALKEVNRQMSYQPDSEISRFNCHTAGKPLRISDFAHVTA EAVHLNRLTH					
	70	80	90	100	110	120
m111.pep	130	140	150	160	170	180
	GALDVTGVLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKIIKQGGKDYASLSKTHPK					
a111	GALDVTGVLVNLWGFDPKSVTREPSPEQIKQAASYTGIDKIIKQGGKDYASLSKTHPK					
	130	140	150	160	170	180
m111.pep	190	200	210	220	230	240
	AYLDLSSIAKGFVDKVGAGELEKYGIQNYLVEIGGELHKGKGNARGEPRWIGIEQPNIVQ					
a111	AYLDLSSIAKGFVDKVGAGELEKYGIQNYLVEIGGELHKGKGNARGEPRWIGIEQPNIVQ					
	190	200	210	220	230	240
m111.pep	250	260	270	280	290	300
	GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNKRPISHNLASISVADSAM					
a111	GGNTQIIVPLNNRSLATSGDYRIFHVDKSGKRLSHIINPNKRPISHNLASISVADSAM					
	250	260	270	280	290	300
m111.pep	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
a111	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLLRX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 399>:

g111-1.seq

```

1 ATGCCGCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTGTC
51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTT GGAACAAACC GCGCAAACCG
101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT
151 TCAAAATATC GGGACAAACT CCCCTCCCTT GCCAAATATC AAAAGCGCAT
201 TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TACCAGACCG

```

```

251 ATTCGGAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ATTTCCGACA CGTTACCGCC GAAGCCGTC GCCTGAACCG
351 CCTGACTCAC GCGCCTCTGG ACGTAACCGT CGGCCCTTTG GTCAACCTTT
401 GGGGGTTCGG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAAACAGG CCGCATCTTA TACGGGCATA GACAAAATCA TTTTGCAACA
501 AGGCAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAATG CGCACGGCGA ACCGTGGCGC ATCGGTATAG
701 AGCAACCCAA TATcatccaa ggcggcaata cgcAGattat cgtccccgtg
751 aaCaaccgtt cgcttgccac ttccggcgAT taccgtaTTT tccacgtcgA
801 TAAAAACGGC Aaacgccttt cccacATCAT CAATCCCAAC AACAAACGAC
851 CCATCAGCCA CAacctcgcc tcCATCagCg TGGTCTCAGA CACTGCAATG
901 ACGGCGGACG GTTTATCCAC ACGATTATTT GTTTTAGGCG AAACCGAAGC
951 CTTAAGGCTG GCAGAACAG AAACACTCGC TGTTTTCTTA ATTGTCGGG
1001 ATAAGGACGG CTACCGCACC GCCATGCTCT CCGAATTGTC CAAGCTGCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 400; ORF 111-1.ng>:

```

g111-1.pep
1 MPSETRLPLNL IRALIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AKIQKRIDDA LKEVNRQMS YQTDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFPGDK SVTREPSPEQ
151 IKQAASYTGI DKILQQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHKG GKNAGEPWR IGIEQPNIVQ GGNTQIIVPL
251 NNRLATSGD YRIFHVDKNG KRLSHIINPN NKRPIHNLA SISVVSADAM
301 TADGLSTGLF VLGETEALRL AEQEKLAFLV IVRDKDGYRT AMSSEFAKLL
351 R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 401>:

```

m111-1.seq
1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGTCT TGATATTGTC
51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTT GGAACAAACC GCGCAAACCG
101 TTACCTTGCA AGGCGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151 TCAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAACGCAT
201 CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGACA CGTTACTGCC GAAGCCGTC GCCTGAACCG
351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTTG GTCAACCTTT
401 GGGGATTCCG CCCCACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAAACAG CCGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
501 AGGCAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGGGCGA ACCGTGGCGC ATCGGTATCG
701 AGCAGCCCAA TATCGTCCAA GCGGCAATA CGCAGATTAT CGTCCCCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
801 TAAAAACGGC AAACGCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCATG
901 ACGGCGGACG GCTTGTCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAAGCTCG TGTTTTCTCT ATTGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGCTCT CCGAATTGTA AAAAGTCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 402; ORF 111-1>:

```

m111-1.pep
1 MPSETRLPLNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMS YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFPGDK SVTREPSPEQ
151 IKQAASYTGI DKILQKQGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHKG GKNARGEPR IGIEQPNIVQ GGNTQIIVPL
251 NNRLATSGD YRIFHVDKNG KRLSHIINPN NKRPIHNLA SISVVSADAM
301 TADGLSTGLF VLGETEALKL AEREKLAFLV IVRDKGGYRT AMSSEFEKLL
351 R*

```

m111-1/g111-1 96.6% identity in 351 aa overlap

```

          10      20      30      40      50      60
m111-1.pep MPSETRLPLNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLNNRDKLPSP
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g111-1      MPSETRLPLNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLNNRDKLPSP

```

326

	10	20	30	40	50	60
m111-1.pep	70	80	90	100	110	120
	AEIQKRIDDALKEVNRQ	STYQPDSEISR	FNQHTAGKPLR	ISSDFAHVTA	EAVRLNRLTH	
g111-1	70	80	90	100	110	120
	AKIQKRIDDALKEVNRQ	STYQTDSEISR	FNQHTAGKPLR	ISSDFAHVTA	EAVRLNRLTH	
m111-1.pep	130	140	150	160	170	180
	GALDVTVGPLVNLWG	FPGDKSVTREPS	PEQIKQAASYT	GIDKII	LKQKDYASLSKTHPK	
g111-1	130	140	150	160	170	180
	GALDVTVGPLVNLWG	FPGDKSVTREPS	PEQIKQAASYT	GIDKII	LKQKDYASLSKTHPK	
m111-1.pep	190	200	210	220	230	240
	AYLDLSSIAKGFV	DKVAGELEKYGI	QNYLVEIGGEL	HGKGNAR	GEPWRIGIEQPNIVQ	
g111-1	190	200	210	220	230	240
	AYLDLSSIAKGFV	DKVAGELEKYGI	QNYLVEIGGEL	HGKGNAR	GEPWRIGIEQPNIVQ	
m111-1.pep	250	260	270	280	290	300
	GGNTQIIIVPLN	NRSLATSGDYR	IFHVDKNGKRL	SHIINPN	NKRPISHNLASISV	VADSAM
g111-1	250	260	270	280	290	300
	GGNTQIIIVPLN	NRSLATSGDYR	IFHVDKNGKRL	SHIINPN	NKRPISHNLASISV	VADSAM
m111-1.pep	310	320	330	340	350	
	TADGLSTGLFVL	GETEALKLAER	EKLAVFLIVR	DKGGYRTAM	SSEFEKLLRX	
g111-1	310	320	330	340	350	
	TADGLSTGLFVL	GETEALKLAER	EKLAVFLIVR	DKGGYRTAM	SSEFEKLLRX	

**g111-1/p44550**

sp|P44550|YOJL\_HAEIN\_HYPOTHETICAL\_LIPOPROTEIN\_HI0172\_PRECURSOR >gi|1074292|pir|C64144  
 hypothetical protein HI0172 - Haemophilus influenzae (strain Rd KW20) >gi|1573128 (U32702)  
 lipoprotein, putative [Haemophilus influenzae Rd] Length = 346  
 Score = 349 bits (885), Expect = 2e-95  
 Identities = 177/328 (53%), Positives = 240/328 (72%), Gaps = 4/328 (1%)

Query: 23 LNACSEQTAQTVTLOGETMGTTYXVKYLSNNRDKLPSPAEIXKRIDDALKEIXNRXMTYQ 82  
 L AC ++T + ++L G+TMGTTY VKYL + S + + I+ LK+ N MSTY+  
 Sbjct: 17 LAACQKET-KVISLSGKTMGTTYHVKYLDGGSITATS-EKTHEEIEAILKDVNAKMSTYK 74

Query: 83 PDSEISRNFQHT-AGKPLRISSDFAHVTA EAVRLNRLTHGALDVTVGPLVNLWGFGPKS 141  
 DSE+SRNFQ+T P+ IS+DFA V AEA+RLN++T GALDVTVG+VNLWGFGP+K  
 Sbjct: 75 KDSELSRNFQNTQVNTPIEISADFAKVLAEAIRLNKVTEGALDVTVGPPVNLWGFGPEKR 134

Query: 142 VTREPSPEQIKQAASYTGIDKII LKQKDYASLSKTHPKAYLDLSSIAKGFVGVKAGEL 201  
 ++P+PEQ+ + ++ GIDKI L K+ A+LSK P+ Y+DLSSIAKGFVGV+VA +L  
 Sbjct: 135 PEKQPTPEQLAERQAWVGIDKITLDTNKEKATLSKALPQVYVDLSSIAKGFVGVQVAEKL 194

Query: 202 EKYGIQNYLVEIGGELHGKGNAR GEPWRIGIEQPNIVQGGNTQIIIVPLNNRSLATSGDY 261  
 E+ QNY+VEIGGE+ KGKN G+PW+I IE+P + ++ LNN +A+SGDY  
 Sbjct: 195 EQLNAQNYMVEIGGEIRAKGNIEGKPWQIAIEKPTTGERAVEAIVGLNMMGMASSGDY 254

Query: 262 RIFHVDKNGKRLSHIINPNKRPISHNLASISVADSAMTADGLSTGLFVLGETEALKLA 321  
 RI+ ++NGKR +H I+P PI H+LASI+V+A ++MTADGLSTGLFVLGE +AL++A  
 Sbjct: 255 RIY-FEENGKRFRAHEIDPKTGYPIQHHLASITVLAPTSMTADGLSTGLFVLGEDKALEVA 313

Query: 322 EREKLAVFLIVRDKGGYRTAMSSEFEKL 349  
 E+ LAV+LI+R G+ T SS F+KL  
 Sbjct: 314 EKNNLAVYLIIRTONGFVTKSSSAFKKL 341

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 403>:

**a111-1.seq**

```

1  ATGCCGCTCTG  AAACACGCCT  GCCGAACCTT  ATCCGCACCT  TGATATTTGC
51  CCTGAGTTT   ATCTTCCTGA  ACGCCTGTTC  GGAACAAACC  GCGCAAACCG
101 TTACCTTGCA  AGGTGAAACG  ATGGGCACGA  CCTATACCGT  CAAATACCTT
151 TCAATAATC   GGGACAAACT  CCCCTCACCT  GCCGAAATAC  AAAAGCGCAT
201 CGATGACGCG  CTTAAAGAAG  TCAACCGGCA  GATGTCCACC  TATCAGCCCG
251 ACTCCGAAAT  CAGCCGGTTC  AACCACACA   CAGCCGC   GCCCCTCCGC
301 ATTTCAAGCG  ACTTCGCACA  CGTTACTGCC  GAAGCGCTCC  ACCTGAACCG

```

327

```

351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCCG CCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAACAAG CAGCATCTTA TACGGGCATA GACAAATCA TTTTGAAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
601 CTGGAATAAT ACGGCATTC AATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGCGCG ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GCGGGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATT TCCACGTCGA
801 TAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
901 ACGGCGGACG GCTTGTCAC AGGATTATTC GTATTGGCGG AACCGGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAAGCTCG TGTTCCTG ATGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTGA AAAAGTGCCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 404; ORF 111-1.a>:

a111-1.pep

```

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMS YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFPGDK SVTREPSPEQ
151 IKQAASYTGI DKILKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPRW IGIEQPNIVQ GGNTQIIVPL
251 NNRLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

a111-1/m111-1 98.9% identity in 351 aa overlap

	10	20	30	40	50	60
a111-1.pep	MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLOGETMGTTYTVKYL	SNNRDKLPSP				
m111-1	MPSETRLPNFIRLVLFALGFIFLNACSEQTAQTVTLOGETMGTTYTVKYL	SNNRDKLPSP				
	70	80	90	100	110	120
a111-1.pep	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTA	EAVHLNRLTH				
m111-1	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISSDFAHVTA	EAVHLNRLTH				
	130	140	150	160	170	180
a111-1.pep	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASYTGIDKIILKQKDY	ASLSKTHPK				
m111-1	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASYTGIDKIILKQKDY	ASLSKTHPK				
	190	200	210	220	230	240
a111-1.pep	AYLDLSSIAKGFVDKVAGELEKYGIQNYLVEIGGELHGKGNARGEPRW	IGIEQPNIVQ				
m111-1	AYLDLSSIAKGFVDKVAGELEKYGIQNYLVEIGGELHGKGNARGEPRW	IGIEQPNIVQ				
	250	260	270	280	290	300
a111-1.pep	GGNTQIIVPLNNRLATSGDYRIFHVDKSGKRLSHIINPNNKRPISHNLA	SISVVADSAM				
m111-1	GGNTQIIVPLNNRLATSGDYRIFHVDKSGKRLSHIINPNNKRPISHNLA	SISVVADSAM				
	310	320	330	340	350	
a111-1.pep	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLL	R				
m111-1	TADGLSTGLFVLGETEALKLAEREKLAVFLIVRDKGGYRTAMSSEFEKLL	R				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 405>:

g114.seq

```

1 ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCAGCAGG AATGCAGCAA
51 GACTTTTTTA TGTCGCGCGG GCGGGACGAG TATGGGGCGG TCAATGTCTGG

```

328

```

101 TAACGGTAGG TTTGTTTTGT GTTTCATTA ACTTAACAAT ATCTGTCGAA
151 TACGGTCAAA GCGGCTATTT TACCAGAGCC GCCGAATGTA AAACAGGGTG
201 TCAGGGCATC AGCCCCGAGCT GCCTGAACGA ACGGACGGTT TGCGAGGTAA
251 CGATAAAATG GTCGAGCAGC GAAACATCAA CCAGCGACAT GGCCTGTGCC
301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAAGGTTTCAG GCGAGCCGCC
351 CGGATGGTTG TCGCGGATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401 GTTTGACGAT TTCGCGGATG TAA

```

This corresponds to the amino acid sequence <SEQ ID 406; ORF 114.ng>:

```

g114.pep
1 MASITSPLHG AQQECSTFL CPPGGTSMGR SMSVTVGLFC VSINLTISVE
51 YGQSGYFTRA AECKTGCGI SPSCLNERTV CEVTIKWSSS ETSTSDMACA
101 SRLVNMSSC EGSGEPPGWL CAIIRLSAYS SNASLTISRM *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 407>:

```

m114.seq
1 ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCACAGAG AATGCAGCAA
51 GACTTTTTTA TGTCCACCGG GCGGGACGAG TATAGGCGCG TCAATGTCGG
101 TAACGGTAGG TTTGTTTTGT GTTTCATTA ACTTAACAAT ATCTGTGAA
151 TACGGTTGAA GCGGCTATTT TATCAGAGCC GCCGCATGTA AAACAGAGTG
201 TCAGGGCATC AACCCGAGCT GTCTGAACGA ACAGACGCTT TGCGAGGTAA
251 CGATAAAATG GTCGAGCAGC GACACATCGA CCAGCGACAT TGCCTGTGCC
301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAAGGTTTCsG GCGAGCCGCC
351 CGGATGGTTG TCGCAATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
401 GTTTGACGAT TTCGCGGATG TAA

```

This corresponds to the amino acid sequence <SEQ ID 408; ORF 114>:

```

m114.pep
1 MASITSPLHG AHRECSKTFI CPPGGTSIGR SMSVTVGLFC VSINLTISVE
51 YGXSGYFIRA AACKTECQGI NPSCLENTL CXVTIKWSSS DTSTSDIACA
101 SRLVNMSSC EXSGEPPGWL CAIIRLSAYS SNASLTISRM *

```

```

m114/g114 90.0% identity over a 140 aa overlap
          10      20      30      40      50      60
m114.pep MASITSPLHGAHRECSKTFICPPGGTSIGRSMSVTVGLFCVSINLTISVEYXSGYFIRA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g114      MASITSPLHGAQQECSTFLCPPGGTSMGRSMSVTVGLFCVSINLTISVEYQSGYFTRA
          10      20      30      40      50      60

          70      80      90      100     110     120
m114.pep AACKTECQGINPSCLENTLCXVTIKWSSSDTSTSDIACASRLVNMSSCEXSGEPPGWL
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g114      AECKTGCGI SPSCLNERTVCEVTIKWSSSETSTSDMACASRLVNMSSCEGSGEPPGWL
          70      80      90      100     110     120

          130     140
m114.pep CAIIRLSAYSSNASLTISRMX
          |||||:|||||:|||||:|||||:|||||
g114      CAIIRLSAYSSNASLTISRMX
          130     140

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 409>:

```

a114.seq
1 ATGCCGGAGG CAAGCATCGC CTCCATCACT TCGCCGCTGC ACGGGGCGCA
51 ACAGGAATGC AGCAAGACTT TTTTATGTCC GCCGGGCGGG ACAGATATGG
101 GGCGGTCAAT GTCGGTAACG GTAGTTTGT TTTGTGTTTC CATTAACTTA
151 ACGATATCTG TCGAATACGG TTGAAGCGGC TATTTTATCA GAGCCGCCGC
201 ATGTAAACA GGGTGTGAGG GCATCAGCCC GAGCTGCCTG AACGAACGGA
251 CGGTTTGC GCATTACGATA AAATGGTCGA GCAGCGACAC ATCGACAGC
301 GACATTGCCT GTGCCAGCCG CCTTGTGAAC ATGATGTCTT CCTGCGAAGG
351 TTCGGGCGAG CCGCCCGGAT GGTGTGCGC GATAATCAGG CTGTCCGCAT
401 ATTCTGCCAA TGCCAGTTTG ACAATTTTAC GGATGTAA

```



This corresponds to the amino acid sequence <SEQ ID 410; ORF 114.a>:

a114.pep

```

1  MPEASIASIT SPLHGAQQEC SKTFLCPPGG TSMGRSMSVT VGLFCVSIINL
51  TISVEYG*SG YFIRAAACKT GCQGISPSCL NERTVCAVTI KWSSSDTSTS
101 DIACASRLVN MMSSCEGSGE PPGWLCAIIR LSAYSSNASL TISRM*

```

m114/a114 92.9% identity in 140 aa overlap

```

              10      20      30      40      50
m114.pep      MASITSPLHGAHRECSKTFLCPPGGTSMGRSMSVTVGLFCVSIINLTISVEYGXSG
              :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a114           MPEASIASITSPHGAQQECSTFLCPPGGTSMGRSMSVTVGLFCVSIINLTISVEYGXSG
              10      20      30      40      50      60

              60      70      80      90     100     110
m114.pep      YFIRAAACKTECQGINPSCLEQTLXVTIKWSSSDTSTSDIACASRLVNMSSCEXSGE
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a114           YFIRAAACKTGCQGISPSCLNERTVCAVTIKWSSSDTSTSDIACASRLVNMSSCEGSGE
              70      80      90      100     110     120

              120     130     140
m114.pep      PPGWLCAIIRLSAYSSNASLTISRMX
              |||||:|||||:|||||:|||||:|||||
a114           PPGWLCAIIRLSAYSSNASLTISRMX
              130     140

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 411>:

g117.seq

```

1  atggtegacg aactegacCT GCTGCCCGAT GCCGTCGCCG CCACCCGTGCT
51  TGCCGACATC GGACGCTACG TCCCCGATTG GAACCTATTG GTTTCCGAGC
101 GCTGCAACAG CACCGTCGCC GAGCTGGTCA AAGGTgtgga CGAAGTGCAG
151 AAACCTACCC ACTTCGCCCG GGTGGACAGC CTCGCCACGC CGGAAGAACG
201 CGCACAGCAA GCGGAAACCA TCGGAAAAT GCTGCTGGCg atgggtaccg
251 Acatcgcggt cgtaTTAATC AAACCTGGCGA TGCGTaccg caccCTGeta
301 ttTTtaagCA ACGCCCCCGA CAGCCCTGAA AAACgcgccg TCgccaaAga
351 aacccTCGAC ATCTTCGCCG CGCTCGCCAA CCGCTTGGGC GTGTGGCAGC
401 TCAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
451 TACCGCGAAA TCGCCTTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
501 ATACATCGAA AACTTCCTCG ATATCCTGCG TACGGAACCT AAAAAATACA
551 ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
601 AAAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGgceTGT TCGACATCCG
651 CGCCGTGCGG ATTCTGGTCG ATACCGTCCC CGaGTGTAC ACCACGCTGG
701 gcaTCGTCCA CAGCCTCTGG CAGCCCATT CCGGCGagtt CGAcgactAC
751 ATCGCCAACC CCAAAGgcaA CGgttATAAA AGtTTGCACA CCGTCATCGT
801 cggcccGGAA gacaaagggt tggaaGtgCA AATCCGCACC TTCGAtatGC
851 accAATTcAa CgaatTcggT gtcgcccGCC ACTGGCGtta caaagaaggc
901 ggcaaaaggcg attccGCcTA cgaacaaAAA ATcgccTggt TGcgcCaACT
951 CTTGGA CTGG CGCGAAAATA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
1001 CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTTATGT TTTGACCCCG
1051 CACGGCAAGG TCCTCTCTCT GCCAACGGGC GCAACCCCCA TCGACTTCGC
1101 CTACGCCCTG CACAGCAGCA TcggCGACCG CTGCCGGGGC GCGAAAGTCG
1151 AaggGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGCGTC
1201 GAAATcatta cCGCcaaAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1251 AGGctgGGtc aAATCCGGCA AGGCCATCGG caaAATCCGC GCCTAcacCC
1301 GCCAGcaaAa cgCgaCACC GTGCGCGAAG AAGGCCGTGT CCAACTCGAC
1351 AAGCAGCTTG CCAAACTCAC GC .AAACCC AACCTGCAAG AGCTTgcga
1401 aaATCTCGCG tacaaAAAGC cagaagacct ctacacCGCc gtcggacaag
1451 gcgaaatttc caaccgcgcc atCcaaaaag cctgcggcac GCTgaacgaa
1501 ccgccccCGG TGCCCGTCAG CGCAACCACC ATCGTCAAAC AGTCCAAAAT

```

```

1551 CAAAAAAGGT GGCCTAACCG GCGTGCTCAT CGACGGCGAA GACGGCTTGA
1601 TGACCACGCT TGCCAAATGC TGCAAACCCG CGCCGCCCGA CGATATTGCC
1651 GGCTTCGTGA CCCGCGAGCG CGGCATTTCG GTCCACCGCA AAACCTGCCC
1701 CTCTTTCCGA CACCTTGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751 GTTGGGCGGC GTTGCGAGAA GGGCAAGTGT TCGCCGTCGA TATCGAAATC
1801 CGC3CCCAAG ACCGCTCCGG GCTTTGCGC GACGTATCCG ACGCGCTCGC
1851 CGCCACAAA CTCAACGTGA CGCCCGTGCA AACCCAGTCC CGCGACTTGG
1901 AAGCCAGCAT GAGGTCACG CTCGAAGTCA AACAAGCAA CGacCTCCCG
1951 CGCGTCCTCG CCGGCTCCG CGATGTCAA GCGTATTGA GCGTATCCCG
2001 GCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 412; ORF 117.ng>:

```

g117.pep
1 MVDELDLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVQ
51 KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVLLI KLAMRTRTLL
101 FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK
151 YREIALLLDE KRTERLEYIE NFLDILRTEL KKYNIHFEVA GRPKHIYSIY
201 KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGFDDY
251 IANPKGNKYK SLHTVIVGPE DKGVEVQIRT FDMHQFNEFG VAAHWRYKEG
301 GKGDSAYEQK IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLTG
351 HGKVLSLPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV
401 EIITAKEGHP SVNWLYEGWV KSGKAIGKIR AYIROQNADT VREEGRVQLD
451 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501 PPPVPVSATT IVKQSKIKKG GKTGVLDIGE DGLMTTLAKC CKPAPPDDIA
551 GFVTRERGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
601 RAQDRSGLLR DVSDALARHK LNVTAQTQS RDLEASMRFT LEVKQVNDLP
651 RVLAGLGDVK GVLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 413>:

```

m117.seq (partial)
1 ..GTGAAACTCA AGAAATACAA TGTCCATTTC GAAGTCGCGC GCGCGCCGAA
51 ACACATCTAC TCCATTTACA AAAAAATGGT GAAGAAAAAA CTCAGCTTCG
101 ACGGCCTCTT TGACATCCGC GCCGTGCGAA TTCTGGTTGA TACCGTCCCC
151 GAGTGTTCAC CCACGCTGGG TATCGTCCAC AGCCTCTGGC AGCCCATTCG
201 CGGCGAGTTC GACGACTACA TCGCCAATCC CAAAGGCAAC GGCATATAAA
251 GTTTGCACAC CGTCATCGTC GGCCCGGAAG ACAAAGGCGT GGAGTACAA
301 ATCCGCACCT TCGATATGCA CCAATTCAAC GAATTCGGTG TCGCCGCCCA
351 CTGgCGTTAC AAAGAGGGCG GCAAGGGCGA TTCCGCTTAC GAACAGAAAA
401 TCGCCTGGTT GCGCCAACTC TTGGACTGGC GCGAAAACAT GCGCGAAAGC
451 GGCAAGGAAG ACCTCGCCGC CGCCTTCAAA ACCGAGCTTT TCAACGACAC
501 GATTATGTT TTGACCCCGC ACGGCAAAGT CCTCTCCCTG CCCACGGGCG
551 CGACCCCAT CGACTTCGCC TACGCCCTGC ACAGCAGCAT CGGCGACCGT
601 TGCCGCGGTG CGAAAGTCGA AGGGCAGATT GTGCCGCTGT CCACCCCGCT
651 CGAAAACGGA CAGCGCGTCG AAATCATTAC CGCCAAAGAA GGCATCCTT
701 CCGTCAACTG GCTTTACGAA GGCTGGGTCA AATCCAACAA GGCAATCGGC
751 AAAATCCGCG CCTACATCCG CCAGCAAAAC GCCGACACCG TCGCGAAGA
801 AGGCCGCGTC CAACTCGACA AACAGCTTGC CAAACTCACG CCCAAACCCA
851 ACCTGCAAGA GCTTGCCGAA AATCTCGGCT ACAAAAAGCC AGAAGACCTC
901 TACACCGCCG TCGGACAAGG CGAAATTTCC AACCGCGCCA TCCAAAAGC
951 CTGCGGCACg CTGAACGAAC CGCCGCCCGT ACCCGTCAGC GAAACCAACA
1001 TCGTCAACAA GTCCAAAATC AAAAAAGGCG GCAAAAACGG CGTGCTCATC
1051 GACGGCGAAG ACGGTCTGAT GACCACGCTT GCCAAATGCT GCAAACCCGC
1101 GCCGCGCGAC GATATTATCG GCTTCGTTAC CCGCGAGCGC GgCATTTCAG
1151 TGCACCGCAA AwyyTkCyCG TCTTTCCAAC ACCTCGCCGA ACACGCGCCC
1201 GAWAAAGTGC TGGACGCAAG CTGGGCGGCA TTGCAGGAAG GACAAGTATT
1251 CGCCGTCGAT ATCGAAATCC GCGCCCAAGA CCGCTCCGGG CTTTTCGCGC
1301 ACGTATCCGA CGCGCTCGCC CGCCACAAAC TCAACGTTAC CGCCGTGCAA
1351 ACCCAGTCCC GCGACTTGGA AGCCAGCATG AGGTTACAGC TCGAAGTCAA
1401 ACAAGTCAAC GACCTCCCGC GCGTCTCGC CAGCCTCGGC GACGTCAAAG
1451 GCGTATTGAG CGTTACCCGG CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 414; ORF 117>:

m117.pep (partial)

331

```

1...VKLKKYNVHF EVAGRPKHIY SIYKKMVKKK LSFGLFDIR AVRILVDTVP
51  ECTTTLGIVH SLWQPIGFE DDYIANPKGN GYKSLHTVIV GPEDKGVEVQ
101 IRTFDMHQFN EFGVAAHWRY KEGGKGDSAY EQKIAWLRQL LDWRENMAES
151 GKEDLAAAFK TELFNDTIYV LTPHGKVLSL PTGATPIDFA YALHSSIGDR
201 CRGAKVEGOI VPLSTPLENG QRVEIITAKE GHPSVNWLYE GWVKSNAIG
251 KIRAYIRQQN ADTVREGRV QLDKQLAKLT PKPNLQELAE NLGYKKPEDL
301 YTAVGQGEIS NRAIQKACGT LNEPPVPVPS ETTIVKQSKI KKGKNGVLI
351 DGEDGLMTTL AKCKPAPPD DIIGFVTRER GISVHRKXXX SFQHLAEHAP
401 XKVLDASWAA LQEGQVFAVD IEIRAQDRSG LLRDVSDALA RHKLNVTAQV
451 TQSRDLEASM RFTLEVQVN DLPRVLASLG DVKGVLVTR L*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 117 shows 97.6% identity over a 490 aa overlap with a predicted ORF (ORF 117.ng) from *N. gonorrhoeae*:

m117/g117

m117.pep				10	20	30
				VKLKKYNVHF	EVAGRPKHIY	SIYKKMVKKK
g117				:	:	:
	150	160	170	180	190	200
m117.pep		40	50	60	70	80
		SFDGLFDIR	AVRILVDTV	PECTTTLG	IVHSLWQPI	PGFE
g117		210	220	230	240	250
m117.pep		100	110	120	130	140
		PEDKGVEVQ	IRT	FDMHQF	NEFGVAAH	WRYKEGGK
g117		270	280	290	300	310
m117.pep		160	170	180	190	200
		KEDLAAAFK	TELFNDTI	YVLT	PHGKVL	SLPTGATP
g117		330	340	350	360	370
m117.pep		220	230	240	250	260
		PLSTPLENG	QRVEIIT	AKEGHPSV	NWLYEGW	VKSNAIG
g117		390	400	410	420	430
m117.pep		280	290	300	310	320
		LDKQLAKLT	PKPNLQEL	AEENLGYK	KPEDLYT	AVGQGEIS
g117		450	460	470	480	490
m117.pep		340	350	360	370	380
		TTIVKQSKI	KKGKNGV	LIDGEDGL	MTTLAKC	CKPAPPDD
g117		510	520	530	540	550
m117.pep		400	410	420	430	440
		FQHLAEHAP	XXVLDASW	AALQEGQV	FAVDIEI	RAQDRSG

[illegible]

a117.seq

1	ATGTTTCATG	AACTCGACCT	GCTCCCCGAT	GCCGTCGCCG	CCACCCTGCT
51	TGCCGACATC	GGACGCTACG	TCCCCGACTG	GAACCTATTG	GTTTCCGAAC
101	GCTGCAACAG	TACCGTCGCC	GAGCTGGTCA	AAGGTGTGGA	CGAAGTGCAG
151	AAACTCACCC	ACTTCGCCCG	GGTGGAACAG	CTCGCCACGC	CGGAAGAACG
201	CGCCGACGAG	CGAGAAACTA	TGCGGAAAAA	GCTGCTGGCG	ATGGTTACCG
251	ACATCCGCGT	CGTGTTAATC	AAACTGGCGA	TGCGTACGCG	CACCCCTGCA
301	TTTTTAAGCA	ACGCCCCCGA	CAGCCCCGAA	AAACGCGCCG	TCGCCAAAGA
351	AACCCCTCGAC	ATCTTCGCCC	CGCTCGCCAA	CCGTTTGGGC	GTGTGGGACG
401	TCAAATGGCA	GCTCGAAGAT	TTGGGCTTCC	GCCATCAAGA	ACCCGAAAAA
451	TACCGCGAAA	TCGCCCTGCT	TTTGGACGAA	AAACGCAACG	AAACGCTCGA
501	ATACATCGAA	AACTTCCTTA	ATATCCTGCG	TACGGAACTC	AAAAAATACA
551	ATATCCACTT	TGAAGTCGCC	GGCCGTCGCA	AACACATCTA	CTCCATTTAC
601	AAAAAAATGG	TGAAGAAAAA	ACTCAGCTTC	GACGGGTGTG	TCGACATCCG
651	CGCGGTGCGG	ATTCTGGTTG	ATACCGTCCC	CGAGTGTTAC	ACCACACTGG
701	GCATTGTCCA	CAGCCTCTGG	CAGCCCATTG	CCGCGGAGTT	GACGACTACG
751	ATCGCCAACC	CGAAAGGCAG	CGGCTATAAA	AGTTTGCACA	CCGTCATCGT
801	CGGCCCGGAA	GACAAAGGCG	TGGAAGTGCA	AATCCGCACC	TTCGATATGC
851	ACCAATTCAA	CGAATTCCGT	GTGCGCCGCG	ACTCGCGTTA	CAAAAGAGGGG
901	GGCAAAGGCG	ATTCCGCTTA	CGACCAAAAA	ATCGCCTGTT	TACGCCAACT
951	TTTGGACTGG	CGCGAAAAAC	TGGCGGAAAG	CGGCAAGGAA	GACCTCGCCG
1001	CCGCCTTCAA	AACCGAGCTT	TTCAACGACA	CGATTTATGT	TTTGACCCCG
1051	CACGGCAAAG	TCCTTCTCCT	GCCCCACAGG	GCGACCCCCA	TCGATTTCCG
1101	CTACGCCCTG	CACAGCAGCA	TCCGCGACCG	TTGCCCGGTT	GCGAAAGTCG
1151	AAGGGCAGAT	TGTGCCGCTG	TCCACCCCCG	TCGAAAAACG	ACAGCGGTGC
1201	GAAATCATTA	CCGCCAAAGA	AGGGCATCCT	TCCGTCAACT	GGCTTTACGA
1251	AGGCTGGGTC	AAATCCAAAC	AGGCAATCGG	CAAAATCCGC	GCCTACATCC
1301	GCGAGCAAAA	CGCCGACACC	GTCGCGCAAG	AAGGCGCCGT	CCAACCTCGAC
1351	AAACAGCTTG	CCAAACTCAC	GCCCCAAACC	AACCTGCAAG	AGCTTGCCCA
1401	AAATCTCGGC	TACAAAAAGC	CAGAAGACCT	CTACACCGCC	CTCGGACAAG
1451	GCGAAATTTC	CAACCGCGCC	ATCCAAAAAG	CCTGCGGCAC	GCTGAACGAA
1501	CCGCCGCCCG	TACCCGTCAG	CGAAACCACG	ATCGTCAAAC	AGTCCAAAAA
1551	CAAAAAAGGC	GGCCAAAAACG	CGGTGCTCAT	CGACGGCGAA	GACGGTCTGA
1601	TGACCACGCT	TGCCAAATGC	TGCAAAACCCG	CGCCGCCCGA	CGACATTGTC
1651	GGCTTCGTTA	CCCGCGATCG	CGGCATTTCG	GTACACCGCA	AAACCTGCCC
1701	CTCTTTCGGA	CACCTCGCCG	AACACGCGCT	CGAAAAAGTA	CTGGACGCAA
1751	GTTGGGCGCG	GTTGCAGGAA	GGACAAGTGT	TCGCCGTCGA	TATCGAAATC
1801	CGCGCCCAAG	ACCCTGCCGG	GCTTTTGCGC	GAGCTATCCG	ACGCGCTTCG
1851	CCGCCACAAA	CTCAACGTTA	CCGCCGTGCA	AACCCAGTCC	CGCGCACTGG
1901	AAGCCAGAGT	GAGGTTACAG	CTCGAAGTCA	AACAAGTTAC	CGACCTCCCA
1951	CGCGTCTCTG	CCAGCCTCGG	CGACGTCAAA	GGCGTATTGA	CGGTTACCCG
2001	GCTTTAA				

a117.pap

1	MVHELDLLPD	AVAATLLADI	GRYPVDWNL	VSERCNSTVA	ELVKGVDEVO
51	KLTHFARVDS	LATPEERAQD	AETMRKMLLA	<u>MVTDIRVVL</u>	<u>KLAMRTRTLQ</u>
101	FLSNAPDSPE	KRAVAKETQD	IFAPLANRLG	VWLQKQLED	LGFRHQEPEK
151	YREIALLLDE	KRTERLEYIE	NFLNLRLTE	KVYNHFEVA	GRPKHYSIY
201	KKMVKKKLSF	DGLFDIRAVR	ILVDTVPECY	TTLGIVHSLV	QPIPGFDDY
251	IANPKNGYK	SLHTVYVGE	DKGVEVQIRT	FDMHQNFEG	VAAHWRVKEG
301	KGKDSAYEQK	IAWLRQLLDW	RHNSAESGKE	DLAAAFKELE	FNDTIYVLTV
351	HGKVLVSLPTG	ATPIDFAYAL	HSSIGDRCRG	AKAEGOVPL	STPLNGORV

333

```

401 EIIITAKEGHP SVNWLYEGWV KSNKAIGKIR AYIRQQNADT VREEGRVQLD
451 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501 PPPVPVSETT IVKQSKIKGK GKNGVLIDGE DGLMTTLAKC CKPAPPDDIV
551 GFVTRDRGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
601 RAQDRSGLLR DVSDALARHK LNVTAQTQS RDLEASMRFT LEVKQVTDLP
651 RVLASLGDKV GVLVSVTRL*

```

m117/a117 98.0% identity in 490 aa overlap

```

                                10      20      30
m117.pep                      VKLNKYNVHFEVAGRPKHIYSIYKKMVKKKL
                                : : : : : : : : : : : : : : : :
a117      EKYREIALLLDLKRTTERLEYIENFLNILRTELAKYNIHFEVAGRPKHIYSIYKKMVKKKL
150      160      170      180      190      200

                                40      50      60      70      80      90
m117.pep      SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKNGYKSLHTVIVG
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKNGYKSLHTVIVG
210      220      230      240      250      260

                                100     110     120     130     140     150
m117.pep      PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSEYEQIAWLRLQLLDWRENMAESG
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSEYEQIAWLRLQLLDWRENMAESG
270      280      290      300      310      320

                                160     170     180     190     200     210
m117.pep      KEDLAAAFKTELFNDTIYVLTTPHGKVLSTPGATPIDFAYALHSSIGDRCRGAKVEGQIV
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      KEDLAAAFKTELFNDTIYVLTTPHGKVLSTPGATPIDFAYALHSSIGDRCRGAKVEGQIV
330      340      350      360      370      380

                                220     230     240     250     260     270
m117.pep      PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQ
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQ
390      400      410      420      430      440

                                280     290     300     310     320     330
m117.pep      LDKQLAKLTPKPNLQELAENLGKYPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSE
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      LDKQLAKLTPKPNLQELAENLGKYPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSE
450      460      470      480      490      500

                                340     350     360     370     380     390
m117.pep      TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKXXXS
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIVGFVTRDRGISVHRKTCPS
510      520      530      540      550      560

                                400     410     420     430     440     450
m117.pep      FQHLAEHAPXKVLDAWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      FRHLAEHAPEKVLDAWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
570      580      590      600      610      620

                                460     470     480     490
m117.pep      QSRDLEASMRFTLEVKQVNDLPRVLASLGDKVGVLSVTRLX
              : : : : : : : : : : : : : : : : : : : : : : : : : : : :
a117      QSRDLEASMRFTLEVKQVTDLPRVLASLGDKVGVLSVTRLX
630      640      650      660

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 417>:

g117-1.seq  
 1 ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CCCTGCAAGA  
 51 ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCGG GACAACGATA  
 101 AAAACCTCAT CGGTACCGCA TGGTCGCTGG CGCAGGAACA TTATCCTGCC  
 151 GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGGCG  
 201 GCGCGAAATG GTCGACGAAC TCGACCTGCT GCCCGATGCC GTCGCCGCCA  
 251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGATTGGAA CCTATTGGTT  
 301 TCCGAGCGCT GCAACAGCAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA  
 351 AGTGCAGAAA CTTACCCACT TCGCCCGGGT GGACAGCCTC GCCACCGCGG  
 401 AAGAACGCGC ACAGCAAGCG GAAACCATGC GGAAATGCT GCTGGCGATG  
 451 GTTACCGACA TCCGCGTCGT ATTAATCAAA CTGGCGATGC GTACGCGCAC  
 501 CTTGCAATTT TTAAGCAACG CCCCCGACAG CCTGAAAAA CGCGCCGCTG  
 551 CCAAAGAAAC CCTCGACATC TTCGCCCGCG TCGCCACCG CTTGGCGGTG  
 601 TGGCAGCTCA AATGGCAGCT CGAAGATTG GGTTCGCGC ATCAAGAACC  
 651 CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC  
 701 CCTCGAATA CATCGAAAA TCCCTCGATA TCCTGCGTAC GGAACCTCAA  
 751 AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC  
 801 CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCGTGTTG  
 851 ACATCCGCGC CGTGGGATT CTGGTCGATA CGGTCCCGCA GTGTTACACC  
 901 ACGCTGGGCA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGagtccGA  
 951 cgactACATC GCCAACCCCA AAGgcaACGg ttATAAAAGt TTGCACACCG  
 1001 TCATCGTcgg cccGGAagaa aaaggtgtgg aagtgcAAAT CCGCACCTTC  
 1051 GATATGCacc AATTCaCga ATTGGGTGTC GCCGCCCACT GCGGTTACAA  
 1101 AGAAGGCGGC AAAGGCGATT CCGCTACGA AAAAAAATC GCTTGGTTGC  
 1151 GCCAACTCTT GGACTGGCGC GAAATATGG CGGAAAGCGG CAAGGAAGAC  
 1201 CTCGCCGCGC CTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT  
 1251 GACCCCGCAC GGCAAAGTCC TCTCTCTGCC AACGGGCGCA ACCCCCATCG  
 1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGCTG CCGGGGCGCG  
 1351 AAAGTCGAAG GGCAGATTGT GCCGTGTCC ACCCCGCTCG AAAACGGACA  
 1401 GCGCGTCGAA ATCATTACCG CCAAGAAGG GCATCCTTCC GTCAACTGGC  
 1451 TTTACGAAGG CTGGGTCAA TCCGGCAAGG CCATCGGCAA AATCCGCGCC  
 1501 TACATCCGCC AGCAAAACGC CGACACCGTG CCGGAAGAAG GCGGTGTCCA  
 1551 ACTCGACAAG CAGCTTGCCA AACTCAGGCC CAAACCCAAC CTGCAAGAGC  
 1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCGCTC  
 1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT  
 1701 GAACGAACCG CGGCCGTGTC CCGTCAGCGC AACCACCATC GTCAAAACAGT  
 1751 CCAAAATCAA AAAAGGTGGC AAAACCGGCG TGCTCATCGA CGGCGAAGAC  
 1801 GGCTTGATGA CCACGCTTGC CAAATGTGTC AAACCCGCGC CGCCCGACGA  
 1851 TATTGCCGGC TTCGTTACCC GCGAGCGCGG CATTTCCGTC CACCGCAAAA  
 1901 CCGTCCCTTC TTTCCGACAC CTTGCCGAAC ACGCGCCCGA AAAAGTACTG  
 1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGG CAAGTGTTTC CCGTCGATAT  
 2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG  
 2051 CGCTCGCCCG CCACAACTC AACGTTACCG CCGTGCAAA CCAATCCCGC  
 2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA  
 2151 CCTCCGCGC GTCCTCGCGG GCCTCGGCGA TGTCAAAGGC GTATTGAGCG  
 2201 TTACCCGGCT TTAA

This corresponds to the amino acid sequence <SEQ ID 418; ORF 117-1.ng>:

g117-1.pep  
 1 MTAISPIQDT QSATLQELRE WFDSCAALP DNDKNLIGTA WSLAQEHYPA  
 51 DAATPYGEPL PDHFLGAAQM VDELDLLPDA VAATLLADIG RYVPDWNLLV  
 101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM  
 151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV  
 201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLDILRTELK  
 251 KYNHFEVAG RPKHIYSIYK KMVKKLSFD GLFDIRAVRI LVDTVPECYT  
 301 TLGIVHSLWQ PIPGEFDDYI ANPKNGYKS LHTVIVGPPE KGEVQIRTF  
 351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED  
 401 LAAAFKEITELF NDTIYVLTPIH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA  
 451 KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SGKAIGKIRA  
 501 YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV  
 551 GQGEISNRAI QKACGTLNEP PPVPVSATTI VKOSKIKKGG KTGVLIDGED  
 601 GLMTTLAKCC KPAPPDDIAG FVTRERGISV HRKTCPSFRH LAEHAPEKVL  
 651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR  
 701 DLEASMRFTL EVKQVNDLPR VLAGLGDVKG VLSVTRL\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 419>:

m117-1.seq  
 1 ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA

```

51  ATTGCGCGAA  TGGTTCGACA  GCTACTGCGC  CGCTCTGCCG  GACAACGATA
101  AAAACCTCAT  CGGTACCGCA  TGGTTGCTGG  CGCAGGAACA  TTACCCCGCC
151  GATGCCGCCA  CGCCGTATGG  CGAGCCGCTG  CCCGACCACT  TCCTCGGCGC
201  GGCAGCAATG  GTTCATGAAC  TCGACCTGCT  CCCCAGATGCC  GTCGCCGCCA
251  CCCTGCTTGC  CGACATCGGA  CGCTACGTCC  CCGACTGGAA  CCTATTGTT
301  TCCGAACGCT  GCAACAGTAC  CGTCGCCGAG  CTGGTCAAG  GTGTGGACSA
351  AGTGCAGAAA  CTCACCCACT  TCGCCCGGGT  GGACAGCCTC  GCCACGCCGG
401  AAGAACCGCG  CCAGCAGGCA  GAAACTATGC  GGAAAATGCT  GCTGGCGATG
451  GTTACCGACA  TCCGCGTCGT  GTTAATCAAA  CTGGCGATGC  GTACGCGCAC
501  CCTGCAATTT  TTAAGCAACG  CCCCAGACAG  CCCCAGAAAA  CGCGCCGTGG
551  CCAAAGAAAC  CCTCGACATC  TTCGCCCGCG  TCGCCAACCG  TTTGGGCGTG
601  TGGCAGCTCA  AATGGCAGCT  CGAAGATTTC  GGCTTCCGCC  ATCAAAAGCC
651  CGAAAAATAC  CGCGAAATCG  CGCTGCTTTT  GGACGAAAAA  CTACCCGAAC
701  CCCTCGAATA  CATCGAAAAA  TTCTCAACA  TCCTGCGCGG  TGAACCTAAG
751  AAATACAATG  TCCATTTCGA  AGTCGCCGGC  CGCCGGAAC  ACATCTACTC
801  CATTTACAAA  AAAATGGTGA  AGAAAAAAT  CAGCTTCGAC  GGCCTCTTTG
851  ACATCCGCGC  CGTGCGAATT  CTGGTTGATA  CCGTCCCCGA  GTGTTACACC
901  ACGCTGGGTA  TCGTCCACAG  CCTCTGGCAG  CCCATTCCCG  GCGAGTTCGA
951  CGACTACATC  GCCAATCCCA  AAGGCAACGG  CTATAAAAGT  TTGCACACCG
1001  TCATCGTCGG  CCGGAAGAC  AAAGGCGTGG  AAGTACAAAT  CCGCACCTTC
1051  GATATGCACC  AATTCACGA  ATTCGGTGTG  GCCGCCCACT  GGCCTTACAA
1101  AGAGGGCGCG  AAGGGCGATT  CCGCCTACGA  ACAGAAAATC  GCCTGGTTGC
1151  GCCAACTCTT  GGACTGGCGC  GAAAACATGG  CGGAAAGCGG  CAAGGAAGAC
1201  CTCGCCGCGC  CCTTCAAAAC  CGAGCTTTTC  AACGACAGCA  TTTATGTTT
1251  GACCCCGCAC  GGCAAAAGTC  TCTCCCTGCC  CACGGGCGCG  ACCCCATCG
1301  ACTTCGCTA  CGCCTGCAAC  AGCAGCATCG  GCGACCGTTG  CCGCGGTGCG
1351  AAAGTCGAAG  GGCAGATTGT  GCCGCTGTCC  ACCCCGCTCG  AAAACGGACA
1401  GCGCGTCGAA  ATCATTACCG  CCAAAGAAAG  GCATCCTTCC  GTCAACTGGC
1451  TTTACGAAGG  CTGGGTCAAA  TCCAACAAGG  CAATCGGCAA  AATCCGCGCC
1501  TACATCCGCC  AGCAAAACGC  CGACACCGTG  CGCGAAGAAG  GCCCGCTCCA
1551  ACTCGACAAA  CAGCTTGCCA  AACTCACGCC  CAAACCCAA  CTGCAAGAGC
1601  TGCGCGAAAA  TCTCGGTAC  AAAAAAGCCAG  AAGACCTCTA  CACCGCGCTC
1651  GGACAAGGCG  AAATTTCCAA  CCGCGCCATC  CAAAAAGCCT  GCGGCACGCT
1701  GAACGAACCG  CCGCCCGTAC  CCGTCAGCGA  AACCACCATC  GTCAAACAGT
1751  CCAAAATCAA  AAAAGGCGCG  AAAAAGCGCG  TGCTCATCGA  CGGCGAAGAC
1801  GGTCTGATGA  CCACGCTTGC  CAAATGCTGC  AAACCCGCGC  CGCCCGACGA
1851  TATATCGGC  TCGTTACCC  GCGAGCGCGG  CATTTAGTGC  CACCGCAAAA
1901  TLGIVHSLWQ  PIPGEFDDYI  ANPKGNGYKS  LHTVIVGPED  KGVEVQIRTF
1951  DMHQFNEFGV  AAHWRYKEGG  KGDSAYEQKI  AWRQLLDWR  ENMAESGKED
401  LAAAFKTELF  NDTIYVLTPL  GKVLSTPTGA  TPIDFAYALH  SSIGDRCRGA
451  KVEGQIVPLS  TPLENGORVE  IITAKEGHPS  VNWLYEGWVK  SNKAIGKIRA
501  YIROQNADTV  REEGRVQLDK  QLAKLTPKPN  LQELAENLGY  KKPEDLYTAV
551  GQGEISNRAI  QKACGTLNEP  PPVPVSETTI  VKQSKIKKGG  KNGVLIDGED
601  GLMTTLAKCC  KPAPPDDIIG  FVTRERGISV  HRKTCPSFOH  LAEHAPEKVL
651  DASWAALQEG  QVFAVDIEIR  AQDRSGLLRD  VSDALARHKL  NVTAVQTQSR
701  DLEASMRFTL  EVKQVNDLPR  VLASLGDVKG  VLSVTRL*

```

This corresponds to the amino acid sequence <SEQ ID 420; ORF 117-1>:

```

m117-1.pep
1  MTAISPIQDT QSATLQELRE WFD SYCAALP DNDKNLIGTA WLLAQEHYPA
51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101  SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151  VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201  WQLKWQLEDL GFRHQKPEKY REIALLLDEK RTERLEYIEN FLNILRGELK
251  KYNVHFVAVG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
301  TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
351  DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWRQLLDWR ENMAESGKED
401  LAAAFKTELF NDTIYVLTPL GKVLSTPTGA TPIDFAYALH SSIGDRCRGA
451  KVEGQIVPLS TPLENGORVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501  YIROQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551  GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601  GLMTTLAKCC KPAPPDDIIG FVTRERGISV HRKTCPSFOH LAEHAPEKVL
651  DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701  DLEASMRFTL EVKQVNDLPR VLASLGDVKG VLSVTRL*

```

m117-1/g117-1 98.2% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSATLQELREWFDSYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL					
g117-1	MTAISPIQDTQSATLQELREWFDSYCAALPDNDKNLIGTAWSLAQEHYPADAATPYGEPL					
	10	20	30	40	50	60
	70	80	90	100	110	120

336

```

m117-1.pep PDHFLGAAQMVHEDLLPDVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
g117-1      PDHFLGAAQMVDELDDLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
              70      80      90      100     110     120

              130     140     150     160     170     180
m117-1.pep LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRLQFLSNAPDSPEK
g117-1      LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRLQFLSNAPDSPEK
              130     140     150     160     170     180

              190     200     210     220     230     240
m117-1.pep RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRRHQKPEKYREIALLLDEKRTERLEYIEN
g117-1      RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRRHQEPEKYREIALLLDEKRTERLEYIEN
              190     200     210     220     230     240

              250     260     270     280     290     300
m117-1.pep FLNILRGELKKYNVHFVAGRPKHIYSIYKKMVKKKLSFDGLFDIPAVRILVDTVPECYT
g117-1      FLDILRTELKKYNHFVAGRPKHIYSIYKKMVKKKLSFDGLFDIPAVRILVDTVPECYT
              250     260     270     280     290     300

              310     320     330     340     350     360
m117-1.pep TLGIVHSLWQPIPGFDDYIANPKNGYKSLHTVIVGPEDKGVQIRTFDMHQFNEFGV
g117-1      TLGIVHSLWQPIPGFDDYIANPKNGYKSLHTVIVGPEDKGVQIRTFDMHQFNEFGV
              310     320     330     340     350     360

              370     380     390     400     410     420
m117-1.pep AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTTPH
g117-1      AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTTPH
              370     380     390     400     410     420

              430     440     450     460     470     480
m117-1.pep GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS
g117-1      GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS
              430     440     450     460     470     480

              490     500     510     520     530     540
m117-1.pep VNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
g117-1      VNWLYEGWVKSGKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
              490     500     510     520     530     540

              550     560     570     580     590     600
m117-1.pep KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED
g117-1      KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSATTIVKQSKIKKGGKNGVLIDGED
              550     560     570     580     590     600

              610     620     630     640     650     660
m117-1.pep GLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASWAALQEG
g117-1      GLMTTLAKCCKPAPPDDIAGFVTRERGISVHRKTCPSFRHLAEHAPEKVLDASWAALQEG
              610     620     630     640     650     660

              670     680     690     700     710     720
m117-1.pep QVFAVDIEIRAQDRSGLLRDVS DALARHKLNVTA VQTOSRDLEASMRFTLEVQVNDLPR
g117-1      QVFAVDIEIRAQDRSGLLRDVS DALARHKLNVTA VQTOSRDLEASMRFTLEVQVNDLPR
              670     680     690     700     710     720

              730
m117-1.pep VLASLGDVKGVLVSVTRLX
g117-1      VLAGLGDVKGVLVSVTRLX
              730

```

m117-1/RelA



sp|P55133|RELA\_VIBSS GTP PYROPHOSPHOKINASE (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE) (PPGPP SYNTHETASE I) >gi|537617 (U13769) ppGpp synthetase I [Vibrio sp.] Length = 744  
Score = 536 bits (1366), Expect = e-151  
Identities = 288/685 (42%), Positives = 432/685 (63%), Gaps = 31/685 (4%)

Query: 74 LDLLPDAVAATLLADI---GRYPDWNLLVSERCNSTVAELVGVDEVQKLTHFARVDSL 130  
L + D + A LL + G Y D + E + T+ LV+GV+++ ++ ++ S  
Sbjct: 68 LSMADATLIAALLYPLVEGGCYSTD---ALKEEYSGTILHLVQGVQMCALS---QLKST 121

Query: 131 ATPEERAQQAETMRKMLLAMVTDIRVVLIKLA MTRTTLQFLSNAPDSPEKRAVAKETLDI 190  
A +A Q + +R+MLL+MV D R V+IKLA R L+ + + PD -RA A+E +I  
Sbjct: 122 AEETAQAAQVDNIRRMLLSMVDDFRCVVIKLAERICNLREVKDCPDEV-RRAAAQECANI 180

Query: 191 FAPLANRLGVWQLKWQLEDLGHQHQPKEKYREIALLLDEKRTERLEYIENFLNIRGELK 250  
+APLANRLG+ QLKW++ED FR+Q P+ Y++IA L E+R +R +YI +F++ L +K  
Sbjct: 181 YAPLANRLGIGQLKWEIEDYAFRYQHPDTYKQIAQLSERRIDREDYITHFVDDLSDAMK 240

Query: 251 KYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPCEYTTLGIVHSLWQ 310  
N+ EV GRPKHIYSI++KM KK L FD LFD+RAVRI+ + + +CY LG+VH+ ++  
Sbjct: 241 ASNIRAQVQGRPKHIYSIWRKMQKKSLEFDELFDVRAVRIVAEELQDCYAAALGVVHTKYR 300

Query: 311 PIPGEFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTFDMHOFNEFGVAHWRYKEG- 369  
+P EDDY+ANPK NGY+S+HTV++GPE K +E+QIRT MH+ +E GVAHW+YKEG  
Sbjct: 301 HLPKEFDDYVANPKNGYQSIHTVVLGPEGKTIEIQIRTKQMHESELGVAHWKYREGT 360

Query: 370 --GKGDSAYEQNIAWLRLQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPHGKVLSP 427  
G SAY++KI WLR+LL W+E M++SG ++ +++++F+D +Y TP G V+ LP  
Sbjct: 361 ASGGAQSAYDEKINWLRKLLAWQEEMSDSG--EMLDELRSQVFDORVYAFTPKGDVVDLP 418

Query: 428 TGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPNVNWLVE- 486  
+ ATP+DFAY +HS +G RC GAKVEG+IVP + L+ G +VEIIT KE +PS +WL  
Sbjct: 419 SNATPLDFAYHINSEVGHRC1GAKVEGRIVPFTYHLQMGDQVEIITQKPNPSRDWLNPN 478

Query: 487 -GWVKSNAIGKIRAYIRQQNADTVREEGRVOLDKQLAKL--TPKPNLQELAENLGKKP 543  
G+V S++A K+ A+ R+Q+ D G+ L+ +L K+ T K A+ K P  
Sbjct: 479 LGFVTSSRARAKVHAWFRKQDRDKNIIAGKEILEAELVKIHATLKDAQYAAKRFNVKSP 538

Query: 544 EDLYTAVGQGEIS-NRAIQACGTLNEPPVPVSETTIVKQSKI-----KKGKNGV 594  
E+LY +G G++ N+ I +N+P + + K S+ KK +- V  
Sbjct: 539 EELYAGIGSGDLRINQVINHINALVNKPTAEEDDQQLLEKLSEASNQATSHKKPQRDAV 598

Query: 595 LIDGEDGLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFOHLAEHAPEKVLDAW 654  
+++G D LMT LA+CC+P P DDI GFVT+ RGISVHR C + L HAPE+++D W  
Sbjct: 599 VVEGVONLMTHLARCCQPIPGDDIQGFVTQGRGISVHRMDCEQLEELRRHAPERIIDTW 658

Query: 655 AALQGEQVFAVDIEIRAQDRSGLLRDVSALARHKLNVAVQTQ--SRDLEASMRFTLEV 712  
G + + + + A +R+GLL+++++ L K+ V +++ + + M F LE+  
Sbjct: 659 GGGFVGN-YTITVRVTASERNGLLKELTNLTMNEKVKVAGMKSRVDYKKQMSIMDFEEL 717

Query: 713 KQVNDLPRVLASLGDVKGVLSVTRL 737  
+ L RVL + VK V RL  
Sbjct: 718 TDLEVLGRVLKRIEQVKDVAEAKRL 742

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 421>:

**a117-1.seq**

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCAC CGCGCTGCCG AACAAACGATA
101 AAAAAGTGTG CTTAGCCGCC CGTTCGCTGG CGGAAGCACA TTACCCCGCC
151 GATGCCGCCA CGCCGTATGG CGAACCCTGG CCCGACCACT TCCTCGGCGC
201 GCGCGAAATG GTTCATGAAC TCGACCTGCT CCCGATGCC GTCGCGGCCA
251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGIT
301 TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351 AGTCGAGAAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCGCG
401 AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAATGCT GCTGGCGATG
451 GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501 CCTGCAATTT TTAAGCAACG CCCCAGACAG CCCCAGAAAA CGCGCCGTCG
551 CCAAAGAAAC CCTCGACATC TTCGCCCGCG TCGCCAACCG TTGGGCGGTG
601 TGGCAGCTCA AATGGCAGCT CGAAGATTG G G .TTCGCC ATCAAGAACC
651 CGAAAAATAC CGCGAAATCG CCTGCTTTT GGACGAAAAA CGCACCGAAC
701 GCCTCGAATA CATCGAAAAC TTCCTTAATA TCCTGCGTAC GGAACTCAAA
751 AAATACAATA TCCACTTTGA AGTCGCCGCG CGTCCGAAAC ACATCTACTC
801 CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGGTTGTTTCG

```

```

851 ACATCCGGCG CGTGGCGATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
901 ACACCTGGGCA TTGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
951 CGACTACATC GCCAACCCGA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCCG CCGGGAAGAC AAAGGCGTGG AAGTGCAAAT CCGCACCTTC
1051 GATATGCACC AATTCAACGA ATTCGGTGTC GCGCGCACT GGCCTTACAA
1101 AGAGGGCGGG AAAGGCGATT CCGCCTACGA AAAAAAATC GCCTGGTTAC
1151 GCCAACTTTT GGA CTGGCGG GAAACATGG CGGAAGCGG CAAGGAAGAC
1201 CTCGCCGCGG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTT
1251 GACCCCGCAC GGCAAGTCC TCTCCCTGCC CACAGGCGCG ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCGCTCG AAAACGGACA
1401 GACGTGTCGAA ATCATTACCG CCAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTACGAAGG CTGGGTCAAA TCCAACAAG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CCGCAAGAAG GCCCTTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCACGCC CAACCCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCGAG AAGACCTCTA CACCGCGGTC
1651 GGACAAGGCG AATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCCCAATC GTCAAAACAGT
1751 CCAAAATCAA AAAAGGCGGG AAAAAGCGCG TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 CATTTGTCGGC TTCGTTACCC GCGATCGCGG CATTTGCGTA CACCGCAAAA
1901 CCTGCCCTCT TTCCGCACAC CTCGCCGAAC ACGCGCCGCA AAAAGTACTG
1951 GACGCAAGTT GGGCGCGGTT GCAGGAAGGA CAAGTGTTTC CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAAC CCAGTCCCGC
2101 GACTTGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAAC AAGTTACCGA
2151 CCTCCACACG GTCCTCGCCA GCCTCGGCGA CGTCAAAAGC GTATTGAGCG
2201 TTACCCGGCT TTAA

```

This corresponds to the amino acid sequence <SEQ ID 422; ORF 117-1.a>:

**a117-1.pep**

```

1  MTAISPIQDT QSATLQELRE WFD SYCTALP NNDKKLVLA RSLAEAHYPA
51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDVQK LTHFARVDSL ATPERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLNLRTELK
251 KYNIHFEVAG RPKHIYSIYK KMKKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKNGYKLS LHTVIVGPED KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWRQLLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGORVE IITAKEGHP S VNWLYEGWK SNKAIGKIRA
501 YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601 GLMTTLAKCC KPAPPDDIVG FVTRDRGISV HRKTCPSFRH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVTDLPR VLASLGDKV VLSVTRL*

```

**a117-1/m117-1** 97.7% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSATLQELREWFD SYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL					
a117-1	MTAISPIQDTQSATLQELREWFD SYCTALPNNDKKLVLAARS LAE AHYPADAATPYGEPL					
	10	20	30	40	50	60
m117-1.pep	PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLV SERCNSTVAELVKGVDEVQK					
a117-1	PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLV SERCNSTVAELVKGVDEVQK					
	70	80	90	100	110	120
m117-1.pep	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIK LAMRTRTLQFLSNAPDSPEK					
a117-1	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIK LAMRTRTLQFLSNAPDSPEK					
	130	140	150	160	170	180
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN					
	190	200	210	220	230	240
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN					

m117-1.pep	250	260	270	280	290	300
	FLN	ILR	GEL	KKY	NVH	FEV
a117-1	250	260	270	280	290	300
	FLN	ILR	GEL	KKY	NVH	FEV
m117-1.pep	310	320	330	340	350	360
	TLG	IVH	SLW	QPI	PGF	DDY
a117-1	310	320	330	340	350	360
	TLG	IVH	SLW	QPI	PGF	DDY
m117-1.pep	370	380	390	400	410	420
	AAH	WRY	KEG	GK	SD	SAY
a117-1	370	380	390	400	410	420
	AAH	WRY	KEG	GK	SD	SAY
m117-1.pep	430	440	450	460	470	480
	GKV	LSL	PTG	ATP	IDF	AYAL
a117-1	430	440	450	460	470	480
	GKV	LSL	PTG	ATP	IDF	AYAL
m117-1.pep	490	500	510	520	530	540
	VNW	LYE	GWV	KS	NKA	IGK
a117-1	490	500	510	520	530	540
	VNW	LYE	GWV	KS	NKA	IGK
m117-1.pep	550	560	570	580	590	600
	KKP	EDL	YAV	GQ	GEI	SN
a117-1	550	560	570	580	590	600
	KKP	EDL	YAV	GQ	GEI	SN
m117-1.pep	610	620	630	640	650	660
	GLM	TTL	AKC	CKP	APP	DDI
a117-1	610	620	630	640	650	660
	GLM	TTL	AKC	CKP	APP	DDI
m117-1.pep	670	680	690	700	710	720
	QVF	AVD	IEI	RAQ	DRS	GLL
a117-1	670	680	690	700	710	720
	QVF	AVD	IEI	RAQ	DRS	GLL
m117-1.pep	730					
	VLAS	LGD	VK	GV	LS	SV
a117-1	730					
	VLAS	LGD	VK	GV	LS	SV

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 423>:

```

g118.seq
1  ATGTGCGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAAA TCATTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGCTG GAGAATGATT TAATcgaGGT TAGGAGAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCCAAATTG GGAGCTTTT GAAATTAAG
251 CTTCCCCTTG GTTGCCTGAT AGCGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAaCA AAAAATAG

```

This corresponds to the amino acid sequence <SEQ ID 424; ORF 118.ng>:

```

g118.pep
1  MCEFkdFRN IPCfEeYDEN sFIGkWyDDG VWDDEeYwKL ENdLIEVRrk
51  YpYPMDIPRD iViGiGTiID fLMvPNwELf EIKaSPwLPD SVGiHERYER
101 FTTMLRYIFT EKdIVNVRFd YYNKK*

```

	70	80	90	100	110	120
m118.pep	VVIGIGTIIDFLMVPNWKLFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD					
a118	IVIGIGTIIDFLMVPNWELFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD					
	70	80	90	100	110	120
m118.pep	YYNKKX					
a118	YYNKKX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 429>:

```

g120.seq
1  ATGATGAAGA CTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAAGGCTACC CCAATCCGCC GTGCTGCACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGTAT ACAATATCCG
201 TTTCAATCC GCGGTACGG TTGTCCGCAA TACCCTGCAC CCTGCCTACT
251 ATAAAGACAT ACGCAGGGGC AAATGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAGGCC
351 CAAGGCTATG GATTGTGTTCA CGCTTGCCTG GCAGTTGGCG GCAATGACG
401 CGAAACTCCC CCCGGTCTG AAAATCACC ACGCAAAAA ACTTTATTCC
451 GTCGGCGGCC TGAATAAGGC GGTACGGGA AAATACAGCA Taggcggcgt
501 gGAAACCGAA GTCGTCAAAT ATCGGTGCG GCGCGCGAC GATACGGTAA
551 CGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGAcgaCG GCAAAACCTA TACGCTGAAG CTCAAATCGG TGCAGATCAA
651 CGGACAGGCC GCCAAACCGT AA

```

This corresponds to the amino acid sequence <SEQ ID 430; ORF 120.ng>:

```

g120.pep
1  MMKTFKNIFS AAILSAAALPC AYAARLPQSA VLHYSYSGYI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNLTH PAYYKDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DTVTYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 431>:

```

m120.seq
1  ATGATGAAGA CTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGmACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGTAT ACAATATCCG
201 TTTCAAGTCC GCGGTACGG TTGTCCGCAA TACCCTGCAC CCTACCTACT
251 ATAGAGACAT ACGCAGGGGC AAATGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAGGCC
351 CAAGGCTATG GATTGTGTTCA CGCTTGCCTG GCAGTTGGCG GCAATGACG
401 CGAAACTCCC CCCGGGCTG AAAATCACC ACGCAAAAA ACTTTATTCC
451 GTCGGCGGTT TGAATAAGGC GGTACAGGA AAATACAGCA TAGGCGGCGT
501 GGAAACCGAA GTCGTCAAAT ATCGGTGCG GCGCGCGAC GATGCGGTAA
551 TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGACGACG GCAAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
651 CGGCCAGGCA GCCAAACCG

```

This corresponds to the amino acid sequence <SEQ ID 432; ORF 120>:

```

m120.pep
1  MMKTFKNIFS AAILSAAALPC A1AAGLPQSA VLXYSGSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNLTH PTYYRDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP

```

343

	10	20	30	40	50	60
m120.pep	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVITYGKAGESKTEQSPKAM					
a120	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVITYGKAGESKTEQSPKAM					
m120.pep	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGGLNKAGTGKYSIGGVETEYVVKYRVRRGD					
a120	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGGLNKAGTGKYSIGGVETEYVVKYRVRRGD					
m120.pep	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDGKTYTLKLKSVQINGQAAPX					
a120	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDGKTYTLKLKSVQINGQAAPX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 435>:

g121.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51  GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCAGCG CTTTACCCCC TACCCTGACC GGTTCGCGCG CAATGTGCTG
151 GATTTCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATAACGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACTGa cgggatttT TACCGTCggc gacttcCGCA
401 GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCGCCTTT
451 CACGAAGCCC TGTTCGCGCA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGCGCA CCCGCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgctg gacgcaggca
601 cacTGGcagc TGCCTTACGA CAAAacggt gcAAAGcgg cacAAGGCAA
651 catatTGcgg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
701 AACCcacc aaAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
751 gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
801 ttcccgattc accgcgcaaA ccgTttggga cgccgtctca CACCGAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGCG ACACGCGTTT CCTGCAACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATttg
1001 cgtggttggC GGCGTGTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 436; ORF 121.ng>:

g121.pep

```

1  METQLYIGIM SGTSMGADA VLVRMDGGKW LGAEGHAFTP YPDRLRRLKLL
51  DLQDTGTDEL HSRRLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTPGP NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRLTSRF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 437>:

m121.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTAGTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCAGCG CTTTACCCCC TACCCCGGCA GGTACGCGG CCAATTGCTG

```

345

```

          310      320      330      340      350      360
m121.pep  LMADLAECFGTRVSLHSTADLNLDPOQWVEAAFAWLAACWINRIPGSPHKATGASKPCIL
          |||
q121      LMADLAECFGTRVSLHSTAE LNLDPOQWVEAAFAWLAACWINRIPGSPHKATGASKPCIL
          310      320      330      340      350      360

m121.pep  XAGYYYYX
          |||
q121      GAGYYYYX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 439>:

```

a121.seq
1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCG TACCCCGGC GGTACGCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAAGTG CACCGCAGCA GGATGTTGTC
201 GCAAGAAGTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCAGCGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTGCGAC
701 AACCCACCC TAAAGCACG GGGCGGAAC TGTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TCGGACGCT
801 TTCCCGATT ACCGCGCAAA CCGTTTTCGA CGCCGCTCA CACGACGGG
851 CAGATGCCCG TCAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCGCGTTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 440; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRRLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPIL AERTQIFTVG DFRSRDLAAG GQGAPLVPF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHKST GRELFALNWL
251 ETYLDGGENR YDVLRTLRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

**m121/a121** 74.0% identity in 366 aa overlap

```

          10      20      30      40      50      60
m121.pep  METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRRLRRQLLDLQDTGADEL
          |||
a121      METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRRLRRQLLDLQDTGADEL
          10      20      30      40      50      60

          70      80      90      100     110     120
m121.pep  HRSRILSQELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPEHGYISQLADLP
          |||
a121      HRSRILSQELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPEHSYSVQLADLP
          70      80      90      100     110     120

```

347

```

m121-1.pep  METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTYPYGRLLRRQLLDLQDTGADEL
              |||||
g121         METQLYIGIMSGTSMGDADAVLVRMDGGKWLGAEGHAFTYPYDRLLRRKLLDLQDTGTDEL
              10      20      30      40      50      60

              70      80      90      100     110     120
m121-1.pep  HRSRILSQELSRLYAQTAELLCSNLAAPS DITALGCHGQTVRHAP EHGYSIQLADLPLL
              |||||
g121         HRSRILSQELSRLYAQTAELLCSNLAAPS DITALGCHGQTVRHAP EHGYSIQLADLPLL
              70      80      90      100     110     120

              130     140     150     160     170     180
m121-1.pep  AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPDDA
              || |||||
g121         AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPDDA
              130     140     150     160     170     180

              190     200     210     220     230     240
m121-1.pep  PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
              |||||
g121         PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFSAQPHPKST
              190     200     210     220     230     240

              250     260     270     280     290     300
m121-1.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICGGGI RNPV
              |||||
g121         GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICGGGI RNPV
              250     260     270     280     290     300

              310     320     330     340     350     360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
              |||||
g121         LMADLAECFGTRVSLHSTADLNLDPOWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
              310     320     330     340     350     360

m121-1.pep  XAGYYYYX
              |||||
g121         GAGYYYYX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 443>:

```

a121-1.seq
1  ATGGAACAC AGCTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGCGATGCC GTACTGATAC GGATGGACGG CGGCAATGCG CTGGGCGCGG
101 AAGGGCACGC CTTACCCCC TACCCCGGCA GGTACGCGG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAACG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCGGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCGGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCGCGACCTT TCGCGCCGGC GGACAAGGCG CGCGCTCGT CCCCCTTTT
451 CACGAAGCCC TGTTCGCGCA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCACGCGA CCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTCGCAC
701 AACCCACCC TAAAAGCACG GGGCGCGAAC TGTTGCCCT AAATGGGTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGATTG ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGACGCG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACGCGCGAA CTGAACCTCG ATCCGCAATG GSTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGCGC CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 444; ORF 121-1.a>:

```

a121-1.pep
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAEE LLCSNLAAPS DITALGCHGQ
101 TVRHAPESY SVOLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPDDA PAFGFDTGPG NMLMDAWMQA

```

349

```

g122.pep
1  MALLSIRKLH KQYGSVTAIQ SLDLDLEKGE VIVLLGPSGC GKSTLLRCVN
51  GLEPHQGGSI VMDGVGEFGK DVSWQTARQK VGMVFQSNEL FAHMTVIENI
101 FLGPVKEQNR DRAEAEAQAG KLLERVGLLD RKNAYPRELS GGQKQRIAI
151 RALCLNPEVI LLDEITAALD PEMVREVLEV VLELAREGMS MLIVTHEMGF
201 ARKVADRIVF MDKGGIVESS DPETFFSAPK SERARQFLAG MDY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 447>:

```

m122.seq
1  GTTGTTCATGA TTAAATCCG CAATATCCAT AAGACCTTTG GCGAAAACAC
51  TATTTTGC GCATCGATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA
101 TCCTCGGGC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC
151 GCGTTGGAAG TGGCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC
201 GCTGAAATC GATTTTCTA AAAAACAAG CAAACACGAT ATTTTGGCAC
251 TGCGCCGCAA ATCAKGCATG GTGTTTCAAC AATACAAYCT CTTTCCGCAC
301 AAAACCCCT TGGAAACGT AATGGAAGGA CCGGTTGCCG TACAGGGCAA
351 GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG
401 GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG
451 CAGCAGCGCG TCGGCATTGC CCGCGCATTG GCGATTGAGC CTGAACTGAT
501 GCTGTTTAC GAACCGACTT CCGCGCTCGA TCCTGAATTG GTGCAAGATG
551 TTTTGGATmC CATGAAGGAA TTGGCGCAAG AAGGCTGGAC CATGTTTGTC
601 GTTACGCATG AAATCAAGTT CGCCTTAGAA GTGGCAACCA CCGwCGTCGT
651 GATGGACrGC GCGGTTATTG TCGAACAAGG CAGCCCCGAA GATTTGTTTC
701 ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTAAGCCA AATCCAATCT
751 ACCAAGATTT GA

```

This corresponds to the amino acid sequence <SEQ ID 448; ORF 122>:

```

m122.pep
1  VVMIKIRNIH KTFGENTILR GIDL DVCKGQ VVILGPSGS GKTTFLRCLN
51  ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSXM VFQQYNLFPH
101 KTALENVMEG PVAVQGPAA QAREEALKLL EKVGLGDKVD LYPYQLSGGQ
151 QQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVLDMKE LAQEGWMTMV
201 VTHEIKFALE VATTXVMDX GVIVEQGSQ DLFDPKHER TRRFLSQIQS
251 TKI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 122 shows 47.2% identity over a 246 aa overlap with a predicted ORF (ORF 122.ng) from *N. gonorrhoeae*:

```

m122/g122
10 20 30 40 50 60
m122.pep VVMIKIRNIHKTFGENTILRGIDL DVCKGQVVILGPSGS GKTTFLRCLNALEMPEDGQI
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
g122 MALLSIRKLHKQYGSVTAIQSLDLDLEKGEVIVLLGPSGC GKSTLLRCVNGLEPHQGGSI
10 20 30 40 50 60

70 80 90 100 110 120
m122.pep EFDNERPLKIDFSKKPSKHDILALRRKSXMVFQQYNLFPHKTALENVMEGPVAVQGPAA
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
g122 VMDGVGEFGKDVSQTA-----RQKVGVMVFQSNELFAHMTVIENIFLGPVKEQNRDRA
70 80 90 100 110

130 140 150 160 170 180
m122.pep QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQ RVGIARALAIQPELMLFDEPTSALDPEL
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
g122 EAEAQAGKLLERVGLLD RKNAYPRELSGGQKQRIAI RALCLNPEVILLDEITAALDP
120 130 140 150 160 170

190 200 210 220 230 240
m122.pep VQDVLDMKELAQEGWMTMVVTHEIKFALEVATTXVMDXGVIVEQGSQD LFDHPKHER
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
g122 VREVLEVLELAREGMSMLIVTHEMGFARKVADRIVFMDKGGIVESSDPETFFSAPKSER

```



351

```

          |||||
a122      TRRFLSQIQSTKIX
          250

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 451>:

```

g122-1.seq
1   ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACCATTTT
51  GCGCGGCATC GATTTGGATG TGGGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCTTCCGG CTCGGGTAAA ACAACATTTT TCGCTGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGCGC GCGCGTTACG
201 CATTGATTTT TCCAAAAAAA CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAGTCCGG AATGGTATTC CAACAATACA ACCTCTTCCC GCATAAAACC
301 GTGTTGAAA ACGTGATGGA AGGGCCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGATAAAGT GGATTTGTAT CCTACCAGC TTTCGGCGG TCAGCAGCAG
451 CGTGTGCGTA TCGCCCGCGC ACTGGCGATT CAGCCTGAAT TGATGCTGTT
501 TGACGAACCC ACTTCCGCGC TGGACCCCGA GTTGGTGCAA GACGTGTTGG
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCACGCT GGAAGTTGCC ACGAACGTCG TCGTGATGGA
651 CGGCGGCGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701 TCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTGCCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 452; ORF 122-1.ng>:

```

g122-1.pep
1   MIKIRNIHKT FGENTILRGI DLDVKGQVV VILGPSGSGK TTFRLCLNAL
51  EMPEDGQIEF DNARPLRIDF SKKTSKHDIL ALRRKSGMVF QQYNLFPHKT
101 VLENVMEGPV AVQKPAQA REEALKLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDAMKELA REGWTMVVVT
201 HEIKFTLEVA TNVVVMDGGV IVEQGSPEL FDHLKHERTR RFLSQIQSAK
251 I*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 453>:

```

m122-1.seq
1   ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACATTTT
51  GCGCGGCATC GATTTGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCTTCCGG CTCAGGCAA ACAGCGTTTC TCGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAATCAGG CATGGTGTTC CAACAATACA ACCTCTTCCC GCACAAAACC
301 GCCTTGAAA ACSTAATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAACT GCTGGAAAAA GTCGGCTTGG
401 GCGACAAAGT GGATTTGTAT CCTACCAGC TTTCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGCGC ATTGGCGATT CAGCCTGAAC TGATGCTGTT
501 TGACGAACCG ACTTCCGCGC TCGATCCTGA ATTGGTGCAA GATGTTTGG
551 ATACCATGAA GGAATTGGCG CAAGAAGGCT GGACCATGGT TGTCGTTACG
601 CATGAAATCA AGTTCGCCTT AGAAGTGGA ACCACGTCG TCGTGATGGA
651 CGGCGGCGTT ATTGTCAAC AAGGCAGCCC GCAAGATTG TTCGACCACC
701 CCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 454; ORF 122-1>:

```

m122-1.pep
1   MIKIRNIHKT FGENTILRGI DLDVCKQVV VILGPSGSGK TTFRLCLNAL
51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQKPAQA REEALKLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDTMKELA QEGWTMVVVT
201 HEIKFALEVA TTVVVMDGGV IVEQGSQDL FDHPKHERTR RFLSQIQSTK
251 I*

```

m122-1/g122-1 94.8% identity in 251 aa overlap

```

          10      20      30      40      50      60
m122-1.pep MIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF
          |||||
g122-1      MIKIRNIHKTFGENTILRGIDLDVKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF

```

q125.seq

51	ATGTCGGGCA	ATGCCTCCTC	TCCTTCATCT	TCCGCCGCCA	TCGGGCTGGT
101	TTGGTTCGGC	GCGGCGGTAT	CGATTGCCGA	AATCAGCACG	GGTACGCTGC
151	TCGCCCCCTT	GGGCTGGCAG	CGCCCTCTGG	CGGCCCTGCT	TTTGGGTGAT
201	GCCGTCGGCG	GCGCGCTGTT	TTTTGGCGCG	CGGTATATCG	GCGCACTGAA
251	CGGACGCAGC	TCGATGGA	GTGTGCGCT	CTGTTCGGC	AAATGCGGTT
301	CAGTGCTGTT	TTCCGTGGCG	AAATGCTGCG	AACTGGCCCG	CTGGACGCGG
351	GTGATGATTT	ACGTGGGCGC	AacggTCAGC	TCCGCTTTGG	GCAAAGTGT
401	GTGGACGgc	gaATCCTTTG	TCTGGTGGGC	ATTGGCAAAC	GGCGCACTGA
451	TCGTGCTGTG	GCTGTGTTTT	GCGCGACGTA	GAACGGGCGG	GCTGAAAACC
501	GTTCGATGC	TGCTGATGCT	GCTTGCCGTG	TTGTGGTTGA	GCGTCGAATT
551	GTTTCGCTTC	TCGCGACAA	ACGCCGCGCC	CGCCGTTTCA	GACGGCACTGA
601	CCTTCGGAAC	GGCAGTCGAA	CTGTCCGCGC	TCATGCCGCT	TTCTTGGCTG
651	CCGTGGGCG	CCGACTACAC	GCGCCAAGCA	CGCCGCCCGT	TTGCGGCAAC
701	CCTGACGGCA	ACGCTCGCCT	ATACCGTGAC	GGGAGTCTGT	ATGATTCGCT
751	TGGGTTTGGC	GCGCGCTCTG	TTTACCGGAG	AAACCGACGT	GCGGAAATCT
801	CTGTTGGGCG	GCGGCTTGGG	CATAACGGGC	ATTCGTGCAG	TGCTCCTCTC
851	CACCGTTACC	ACAACGTTTC	TCGATACCTA	TTCCGCCGGC	GCGAGTGCGA
901	ACAACATTTC	CGCGCGTTTT	GCGGAAATAC	CCGTGCTGCT	CGCGGTTACC
951	CTGatccgca	cctgtcctgc	cgtcatgctg	cccggtaccg	aatataaaaa
1001	cttcctgctg	cttatccctg	cggatattgg	gccgatggcg	ggtgggtttg
1051	attgcgcaCT	TTTttctctt	AAACGCGCGT	GA	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 459>:

m125.seq

1	ATGTCGGGCA	ATGCCTCCTC	TCCTTCATCT	TCCTCCGCCA	TCGGGCTGAT
51	TTGGTTCGGC	GCGGCGGTAT	CGATTGCCGA	AATCAGCACG	GGTACCGCTG
101	TTGCGCTCTT	GGGCTGGCAG	CGCGGCTCGG	CGGCTCTACT	TTTGGGTCAAC
151	GCCGTCGGGG	GCGGCTGTGT	TTTTGGGGGG	GGGTATATCG	CGCGACTGAC
201	CGGACGCAGC	TCGATGGAAA	CGCTGCGCCT	GTCGTTTCGGC	AAACCGCGTT
251	CAGTGCTGTT	TTCCGTGGCG	AATATGCTGC	AACTGGCCGG	CTGGACGCGC
301	GTGATGATTT	ACGCCGGCGC	AACGGTCAGC	TCGGCTTTGG	GCAAAAGTGT
351	CTGGGACGGC	GAATCTTTTG	TCTGGTGGGC	ATTGGCAAA	GGCCCGCTGT
401	TTGTGCTGTG	GCTGGTTTTC	GGCGCACGCA	AAACAGCGCG	GCTGAAAACC
451	GTTTCGATGC	TGCTGATGCT	GTTGGCGGTT	CTGTGGCTGA	GTGCCGAAGT
501	CTTTTCCACG	GCAGGCCAGCA	CCGCCCGACA	GGTTTTCAGC	GGCATGAGTT
551	TCGGAAACGG	AGTCGAGCTG	TCCGCCCTGA	TGCCGCTTTC	CTGGCTGCCG
601	CTTGCGCGCG	ACTACACGCG	CCACGCGCGC	CGCCGGTTTG	CGGCAACCTT
651	GACGGCAACG	CTGCGCTACA	CGCTGACCGG	CTGCTGGATG	TATGCCTTGG
701	GTTTGGCAGC	GGCGTGTGTC	ACCGGAGAAA	CCGACGTGGC	AAAAATCTCT
751	CTGGGCGCAr	GTTTgGTGTC	GGCAGGCATT	TGGCGGTCG	TCCTCTCCAC
801	CGTTTACCACA	ACGTTTCTCG	ATTGCCTATC	CGCCGGCGCG	AGTGCGAACAC

```

a125.pep
  1  MSGNASSPSS  SAAIGLIWFG  AAVSIAEIST  GTLLAPLGWQ  RGLAALLLGH
  51  AVGGALFFAA  AYGALGTGRS  SMESVRLSFG  KRGSVLFSPA  NMLQLAGWTA
101  VMTYAGATVS  SALGKVLWDG  ESFVWAAALN  GALIVLWLVF  GARKTGGKLT
151  VSMLLMLLAV  LWLSAEVFST  AGSTAAQVSD  GMSFGTAVEL  SAVMPLSWLP
201  LAADYTRHAR  RPFATLTAT  LAYLTGSCWM  YALGLAAALF  TGETDVAKLL
251  LGAGLGAAGI  LAVLSTVTVT  TFLDAYSAGV  SANNISAKLS  EPIPIAVAVV
301  VGTLLAVLLP  VTEYENFLLL  IGSVFAPMAX  GFDCLRFRLR  TA*

```

	10	20	30	40	50	60
m125.pep	MSGNASSPSSSSAIGLWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA					
a125	::					
	10	20	30	40	50	60
	70	80	90	100	110	120
m125.pep	AYIGALTGRSSMESVRLSFGKRGSVLFVANMLQLAGWTAVMIYAGATVSSALGKVLWDG					
a125						
	70	80	90	100	110	120
	130	140	150	160	170	180
m125.pep	ESFVWWALANGALIVLWLWVFSGARKTGGLKTVMMLMLLAVLWLSAEVFTAGSTAAQVSD					
a125						
	130	140	150	160	170	180
	190	200	210	220	230	240
m125.pep	GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCMYALGLAAALF					
a125						
	190	200	210	220	230	240
	250	260	270	280	290	300
m125.pep	TGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVTL					
a125						
	250	260	270	280	290	300
	310	320	330	340		
m125.pep	IGTVLAVMLPVTEYENFLLLIGSVFAPMAGGFDCRRLFRETAX					
a125	: :: ::					
	VGTLLAVLLPVTEYENFLLLIGSVFAPMAGGFDCRRLFRETAX					

357

```

101 DWIKLELIGD DDTLQPDVFQ LVEAAEILIK DGFKVLPYCT EDLIACRRLL
151 DAGCQALMPW AAPIGTGLGA VHAYALNVLR ERLPDTPLII DAGLGLPSQA
201 AQVMWGFDFG VLLNTAVSRS GDPVNMARAF ALAVESGRLA FEAGPVEARD
251 KAQASTPTVG QPFWHSAEY*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 126 shows 95.9% identity over a 269 aa overlap with a predicted ORF (ORF 126.ng) from *N. gonorrhoeae*:

m126/g126

```

                                10      20      30      40
m126.pep                      HYTKPEIMLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQ
                                ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      MPSETPKARRRLSDGIASDNHTKESIMLTLYGETFPSRLLLGTAAYPTPEILKQSVRTAR
                                10      20      30      40      50      60

                                50      60      70      80      90      100
m126.pep      PAMITVSLRRAGSGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      PAMITVSLRRTGCGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETD
                70      80      90      100      110      120

                                110      120      130      140      150      160
m126.pep      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                130      140      150      160      170      180

                                170      180      190      200      210      220
m126.pep      APIGTGLGAVHAYALNVLRERLPDTPLIIDAGLGLPSQAAQVMWGFDFGVLLNTAVSRSG
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      APIGTGLGAVHAYALKILRERLPDTPLIIDAGLGLPSQAAQVMWGFDFGVLLNTAVSRSG
                190      200      210      220      230      240

                                230      240      250      260      270
m126.pep      DPVNMARAFALAVESGRLA FEAGPVEARDKAQASTPTVGQPFWHSAEYX
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      DPVNMARAFALAVESGRLA FEAGPVEARTKAQASTPTVGQPFWHSAEYX
                250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 467>:

```

a126.seq
1   TTGTTAATCC ACTATACAAA GGAACCCATT ATGCTCACCC TGTACAGCGA
51  AACTTTCCTT TCGCGGCTGC TGCTCGGCAC AGCCGCCTAC CCGACCCCTG
101 AAATCCTCAA ACAATCCGTC CGAACC GCCCGCGAT GATTACCGTC
151 TCGCTGCGCC GCGCGGGATG CGGCGGCGAG GCGCACGTC AGGGGTTTGT
201 GTCGCTGCTT CAAGAAACCG GCGTTCCCGT CCTGCCGAAC ACGGCAGGCT
251 GCCAAAGCGT GCAGGAAGCG GTAACGACGG CGCAAATGGC GCGCGAAGTG
301 TTTGAAACCG ATTGGATTAA ACTCGAACTC ATCGGCGACG ACGACACCTT
351 GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC GCGGGAATC CTGATTAAAG
401 ACGGCTTCAA AGTGCTGCCT TATTGCACCG AAGACCTGAT TGCCTGCCGC
451 CGCCTGCTCG ACGCGGCTG TCAGGCGTTG ATGCCGTGGG CGGCCCGAT
501 CGGCACGGGT TTGGGCGCGG TTCACGCCTA CGCGTTGAAC GTCCTGCCGC
551 AACGCCTGCC CGACACGCCG CTTGATTATCG ACGCGGCTT GGGTTTGCCC
601 TCACAGGCGG CACAAGTGAT GGAATGGGGC TTTGACGGCG TGCTTTTGAA
651 TACTGCCGTT TCCCGCAGCG GCGATCCGGT CAATATGGCA CGCGCCTTCG
701 CACTCGCCGT CGAATCCGGA CGGCTGGCAT TTGAAGCCGG ACCGGTCGAA
751 GCACGCGACA AAGCGCAAGC CAGCAGCCG ACAGTCGGAC AACCGTTTTG
801 GCATTCCGCG GAATATTGA

```

359

```

51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVVEAAEI LIKDGFKVLP YCTEDLIACR RLIDAGCQAL
151 MPWAAPIGTG LGAVHAYALK ILRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARTKAQASTP
251 TVGQPFWWSA EY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 471>:

**m126-1.seq**

```

1 ATGCTCACCC TATACGGCGA AACTTTCCCC TCGCGGCTGC TGCTCGGCAC
51 GGCTGCCTAC CCGACCCCGG AAATCCTCAA ACAATCCATC CAAACCGCCC
101 AGCCTGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGAAG CCGCGGCGAG
151 GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAAACGACGG
251 CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAACTC
301 ATCGGAGATG ACGACACCTT GCAGCCGGAT GTGTTCAGC TTGTGCAAGC
351 GCGGGAATC CTGATTAAAG ACGGCTTCAA AGTCTGCTCT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCGGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CCGGTTGAAC GTCCTGCGCG AAGCCTGCCG CGACACGCGG CTGATTATCG
551 ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651 CAATATGGCA CGCGCCTTCG CACTCGCGGT CGAATCCGGA CGGCTGGCAT
701 TTGAAGCCGG ACCGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
751 ACAGTCGGAC AACCGTTTGT GCATTCGGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 5; ORF 126-1>:

**m126-1.pep**

```

1 MLTYGETFP SRLLLGTAAY PTPEILKQSI QTAQPAMITV SLRRAGSGGE
51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVVEAAEI LIKDGFKVLP YCTEDLIACR RLIDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWWSA EY*

```

**m126-1/g126-1** 96.9% identity in 262 aa overlap

	10	20	30	40	50	60
m126-1.pep	MLTYGETFP SRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGEAHGQGFWSLL					
	10	20	30	40	50	60
g126-1	MLTYGETFP SRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRTGCGGEAHGQGFWSLL					
	70	80	90	100	110	120
m126-1.pep	QETGVPVLPNTAGCQSVQEA VTTAQMAREV FETDWIKLELIGDDDTLQPDVFQVVEAAEI					
	70	80	90	100	110	120
g126-1	QETGVPVLPNTAGCQSVQEA VTTAQMAREV FETDWIKLELIGDDDTLQPDVFQVVEAAEI					
	130	140	150	160	170	180
m126-1.pep	LIKDGFKVLPYCTEDLIACRRLDAGCQALMPWAAPIGTGLGAVHAYALNVLRLRPLDTP					
	130	140	150	160	170	180
g126-1	LIKDGFKVLPYCTEDLIACRRLDAGCQALMPWAAPIGTGLGAVHAYALKILRLRPLDTP					
	190	200	210	220	230	240
m126-1.pep	LIIDAGLGLPSQAAQVMEWGFQGVLLNTAVSRSGDPVNMA RAFALAVESGRLAFEAGPVE					
	190	200	210	220	230	240
g126-1	LIIDAGLGLPSQAAQVMEWGFQGVLLNTAVSRSGDPVNMA RAFALAVESGRLAFEAGPVE					
	250	260				
m126-1.pep	ARDKAQASTPTVGQPFWWSAEYX					
	250	260				
g126-1	ARTKAQASTPTVGQPFWWSAEYX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 473>:

**a126-1.seq**

```

1 ATGCTCACCC TGTACAGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
51 AGCCGCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
101 GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGATG CCGCGGCGAG

```

361

```

551 CCGTCGAAAT CCCCGTTCCC ATCCATTGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCGCCCAT
651 TCAGCGGTAT TTGGAACG TGCAGGCGGA AAAACTGTTT ATCAGCCCCG
701 CCGCCAGGCC GCGCGTTACC CCGGTACCGT ACGACGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCTC CCCGTTTCA AAGCGGCTGG AAATCCAACA
801 GGCGGTTATG GACGAATTTT TCGCGTACA ATACCGCCTG TTAATCATC
851 CCGCCGgctc cgAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 476; ORF 127.ng>:

```

g127.pep
1  MEIWNMLNTW PDAVPIRAEA AESVAAVAAL LLARALLNI HFRRHPDFGI
51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAADVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251 IIVRFASPVs KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 477>:

```

m127.seq
1  ATGGAAATAT GGAATATGTT GGACACTTGG CTCGGTGCCG TCCCGATACG
51  TCGCGAGGCG GTCGAATCCG TGGCGGCGGT TCGCGCTTGG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAAAC GGCATCCGGA TTTGCGCATC
151 GAAAGCAAGC GCGCGTTTTT GGTGCGCAGC CGCAATATAA CGCTGCTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GCGCGAAATC CAAACGCTGG
251 CTTTGTTCGAT GTTTGCGGTG GCGGCGGCGG TCGTCTGGGC GACGAAGGAA
301 CTGATTATGT GTCTGTCGGG CAGTATTTTA AGGTCTGCCA CCCAGCAATA
351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401 ACATCAACCT GTTGAACACG CTGATGATGC AGGTCTGGTC GAACCCCTTG
451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCGCGG ACAATATTTT GGGCGACTAT GTCATCCATA
551 CCGTCGAAAT CCCCGTTCCC ATCCATTGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCGCCCAT
651 CCAACGGSAT TTGGAACG TGCAGGCGGA AAAACTGTTT ATCAGCCCCG
701 CCGCCAGACC GCGCGTTACC CCGGTGCGGT ACGATGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCTTC CCCGTTTCA AAGCGGCTGG AAATCCAACA
801 GGCGGTTATG GACGAATTTT TCGCGTACA ATACCGCCTG TTAATCACC
851 CCGCCGCTC CGAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 478; ORF 127>:

```

m127.pep
1  MEIWNMLDTW LGAVPIRAEA VESVAAVAAL LLARALLNI HFKRHPDFGI
51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAADVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251 IIVRFASPVs KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 127 shows 97.9% identity over a 290 aa overlap with a predicted ORF (ORF 127.ng)

from *N. gonorrhoeae*:

```

m127/g127
      10      20      30      40      50      60
m127.pep  MEIWNMLDTWLGAVPIRAEAESVAAVAALLLARALLNIHFKRHPDFGIESKRRFLVAS
          |||||:|  |||||:|  |||||  |||||  |||||  |||||  |||||  |||||
g127      MEIWNMLNTWPDVAVPIRAEAESVAAVAALLLARALLNIHFRRHPDFGIESKRRFLVAS
          10      20      30      40      50      60

      70      80      90      100     110     120
m127.pep  RNITLLLVLFSLAFIWSAQIQTALALSMFAVAAAVVATKELIMCLSGSILRSATQQYSVG
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
g127      RNITLLLVLFSLAFIWSAQIQTALALSMFAVAAAVVATKELIMCLSGSILRSATQQYSVG

```

363

	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGLAGTTVSFPNSLLLSHPVRRDNLGDY					
a127						
	130	140	150	160	170	180
	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGLAGTTVSFPNSLLLSHPVRRDNLGDY					
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLSDSEAVCRKAVLEPLCAPYIPAIQRKLENVQAEKLFITPAARPRVT					
a127						
	190	200	210	220	230	240
	VIHTVEIPVPIHLSDSEAVCRKAVLEPLCAPYIPAIQRHLENVQAEKLFITPAARPRVT					
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFRLVQYRLLNHPAGSETLX					
a127						
	250	260	270	280	290	
	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFRLVQYRLLNHPAGSETLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 481>:

```

g128.seq
1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca
51  aatccaaacc gaagACatca AACCCGCGT CCAAACCGCC ATCGCGGAAG
101 CGCGCGGACA AATCGCGGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCAGCACC
401 TGC CGCATTT CGTATTGACC GCGCGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAA CGAAGGCGCG CAACTTTCCG CCAAATTCCT
501 CCAAACGCTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGCGCAT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTATACC AATACGCGCG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTTCGCCGC GAACACCTCG GTCTCGCGA CCCGACGCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTTCTGGCAG GCCTGTTCGC
1101 CCAAATCAAA AAATCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACA AGACGTGCGC TATTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGCGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACCGCGCGG
1251 CGCGTGGATG AACGACTaca AAGGCCGCGC CCGCTTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AacCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAAGTGGCG
1451 TGTCCGGCAT CAacggcgtA GAATGGGACG CGGTGCAACT GCCCAGCCAG
1501 TTTATGGAAA ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551 CCACGAAGAA AccgGCGAGC CCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TcgCGCCAA AAATTCAG CGCGGTATGT TCCTCGTCCG GCAAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGTCT
1701 GAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGAGG CTATTACAGC TACGCATGGG CCGAAGTCCT
1851 cAGCACCGAT GCCTACGCGC CTTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951 gcgCGGAAT CTTCAAAGC CTTCGCGGA CGCGAACCGA GCATAGACGC

```

301 GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from *N. gonorrhoeae*:

m128/g128

	10	20	30	40	50	60
g128.pep	MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQHTGWANTVERLTGIT					
m128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIAQHTGWANTVEPLTGIT					
	10	20	30	40	50	60
g128.pep	70	80	90	100	110	120
	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA					
m128	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
g128.pep	130	140	150	160	170	180
	TLSPAQKTKLDHDLRDFVLSGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128	TLSPAQKTKLNH					
	130					
	//					
g128.pep				340	350	360
	YAGEKLREAKYAFSETEVKKYFPVGKVLG					
m128	YASEKLREAKYAFSETXVKKYFPVGXVLNG					
	10	20	30			
g128.pep	370	380	390	400	410	420
	LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQONGKTIGGVMDLYAREGKRGGAWMNDYK					
m128	LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQONGEXIGGVMDLYAREGKRGGAWMNDYK					
	40	50	60	70	80	90
g128.pep	430	440	450	460	470	480
	GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVDELGV					
m128	GRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVDELGV					
	100	110	120	130	140	150
g128.pep	490	500	510	520	530	540
	SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQXGMF					
m128	SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKMLAAKNFQXGMF					
	160	170	180	190	200	210
g128.pep	550	560	570	580	590	600
	LVRQMEFALFDMMIYSEDECRLKNWQQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY					
m128	XVRQXEFALFDMMIYSEDDGRLKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY					
	220	230	240	250	260	270
g128.pep	610	620	630	640	650	660
	SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS					
m128	SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS					
	280	290	300	310	320	330



367

551 FALFDMMIYS EDDEGR LKNW QQVLDSVRKE VAVVRPPEYN RFANSEFGHIF  
 601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS  
 651 AAESFKAFRG REPSIDALLR HSGFDNAA\*

m128/a128 66.0% identity in 677 aa overlap

m128.pep	10	20	30	40	50	60
	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT					
a128	MTONALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT					
	10	20	30	40	50	60
m128.pep	70	80	90	100	110	120
	ERVGRIGVGVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
a128	ERVGRIGVGVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
m128.pep	130	TLSPAQKTKLNH-----				
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
m128.pep	-----					
a128	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLRQIYRAYV					
	190	200	210	220	230	240
m128.pep	-----					
a128	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMDTPEQVLNFLHDL					
	250	260	270	280	290	300
m128.pep	-----					140
	-----					150
a128	ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVVGK					YASEKLREAKYAFSETXVKKYFPVVGX
	310	320	330	340	350	360
m128.pep	160	170	180	190	200	210
	VLNGLFAQXKKLYGIGFTEKTPVWHKDVRYXELQONGEXIGGVYMDLYAREGKRGGAWM					
a128	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQONGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
m128.pep	220	230	240	250	260	270
	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD					
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPVPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
m128.pep	280	290	300	310	320	330
	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ					
a128	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540
m128.pep	340	350	360	370	380	390
	XGMFXVRQXEFALFDMMIYSEDDEGR LKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	RGMFLVRQMEFALFDMMIYSEDDEGR LKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					
	550	560	570	580	590	600

```

201 GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCGGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACACGATC
401 TGCGCGATTT CGTCCTCAGC GCGCGGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAACTGGCAA AACTGCAAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTGCCGCGC
601 GCGCGCGAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 ACACTACCTC GCGCTCATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTITC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAACTA CGCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCTC GCACGACCTC
901 GCGCGCGCGC CCAAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAACTG CCGCAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTATTAAACG GACTGTTGCG
1101 CCAAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCGC
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAA CGCGCAAAAC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCGCGCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCCGCTGAG CCACGACGAA ATCCTCATCC TTTTCACGA
1401 AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGCG
1451 TATCGGCAT CAACGGCGTA GAATGGGACG CGGTGCAACT GCCAGCCAG
1501 TTTATGAAA ATTTCGTTT GGAATACAAT GTCTGGCAC AAATGTCAGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAATCTTTC GACAAAATGC
1601 TCGCGGCAA AAATTCCAA CGCGGCATGT TCCTCGTCCG GCAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGAGC AAGGCGGTCT
1701 GAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCGTCA
1751 TCCAGCGGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCGGCA
1901 CAGGCAACG CTTTGGCAG GAAATCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCCGCGG CGCGAACCGA GCATAGACGC
2001 ACTCTGCGC CACAGCGGTT TCGACAAACG GGTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 490; ORF 128-1>:

m128-1.pep.

```

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVPELTGIT ERVGRWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLANALQTA LGGFKNYAE LSLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351 EVKKYFPVVK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGAWM NDYKGRRRFS DGTQLQPTAY LVCNFAPPVG
451 GREARLSHDE ILILFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLLKNW QQVLDVSRKK VAVIQPPEYN RFALSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAEFPKAFRG REPSIDALLR HSGFDNAV*

```

m128-1/g128-1 94.5% identity in 491 aa overlap

```

          10      20      30      40      50      60
g128-1.pep MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQTHTGWANTVERLTGIT
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90     100     110     120
g128-1.pep ERVGRWGVVSHLNSVADTP ELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128-1      ERVGRWGVVSHLNSVADTP ELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90     100     110     120

```

371

```

1251 CCGGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCAGGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACC GGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGGAAA ATTTCTGTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAAGTCTTC GACAAAAATGC
1601 TCGCCGCCAA AACTTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGCGCGCT ATTCGCGCAG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCGC CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 492; ORF 128-1.a>:

a128-1.pep

```

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWGA
51 NTVEPLTGIT ERVGRIGVVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHLY AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVVHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLEKNW QQVLDVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ SLAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*

```

m128-1/a128-1 97.8% identity in 677 aa overlap

	10	20	30	40	50	60
a128-1.pep	MTDNALLHLG	EEPRFDQIKT	EDIKPALQTA	IAEAREQIAA	IKAQHTGWANT	VEPLTGIT
m128-1	MTDNALLHLG	EEPRFDQIKT	EDIKPALQTA	IAEAREQIAA	IKAQHTGWANT	VEPLTGIT
	10	20	30	40	50	60
a128-1.pep	ERVGRIGVVV	SHLNSVTDTP	ELRAAYNELM	PEITVFFTEI	GQDIELYNRF	KTIKNSPEFD
m128-1	ERVGRIGVVV	SHLNSVADTP	ELRAVYNELM	PEITVFFTEI	GQDIELYNRF	KTIKNSPEFD
	70	80	90	100	110	120
a128-1.pep	TLSHAQTKL	NHDLRDFVLS	GAELPPEQQA	ELAKLQTEGA	QLSAKFSQNV	LDATDAFGIY
m128-1	TLSPAQTKL	NHDLRDFVLS	GAELPPEQQA	ELAKLQTEGA	QLSAKFSQNV	LDATDAFGIY
	130	140	150	160	170	180
a128-1.pep	FDDAAPLAGI	PEDALAMFAA	AAQSEGKTGY	KIGLQIPHLY	AVIQYADNRK	LREQIYRAYV
m128-1	FDDAAPLAGI	PEDALAMFAA	AAQSEKTYK	IQLQIPHLY	AVIQYADNREL	LREQIYRAYV
	190	200	210	220	230	240
a128-1.pep	TRASELSDDG	KFDNTANIDR	TLENALQTAK	LLGFKNYAEL	SLATKMADTP	EQVLNFLHDL
m128-1	TRASELSDDG	KFDNTANIDR	TLENALQTAK	LLGFKNYAEL	SLATKMADTP	EQVLNFLHDL
	250	260	270	280	290	300
a128-1.pep	ARRAKPYAEK	DLAEVKAFAR	ESLGLADLQP	WDLGYAGEKL	REAKYAFSET	EVKKYFPVGK
m128-1	ARRAKPYAEK	DLAEVKAFAR	ESLGLADLQP	WDLGYAGEKL	REAKYAFSET	EVKKYFPVGK
	310	320	330	340	350	360
a128-1.pep	ARRAKPYAEK	DLAEVKAFAR	ESLGLADLQP	WDLGYAGEKL	REAKYAFSET	EVKKYFPVGK
m128-1	ARRAKPYAEK	DLAEVKAFAR	ESLGLADLQP	WDLGYAGEKL	REAKYAFSET	EVKKYFPVGK

Query: 363 NGLFAQIKKLYGIGFTE-KTVPVWHKDVRYFEL-QQNGETIGGVYMDLYAREGKRGAWM 420  
 +GLF IK+++ I E K V WHKDVRF+L +N + G Y+DLYARE KRGAWM  
 Sbjct: 365 SGLFELIKRIFNIRAVERRKGVDTWHKDVRFDLIDENDQLRGSFYLDLYAREHKRGAWM 424

Query: 421 NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEIXXXXXXXXXXXXXXQVD 480  
 +D GR+R DG+++ P AYL CNF P+G K A +H+E+ Q+D  
 Sbjct: 425 DDCIGRKRKLDGSIETPVAYLTCNFNAPIGNKPAFTHNEVTTLFHEFGHGIHMLTQID 484

Query: 481 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ 540  
 V+GINGV WDAVELPSQFMEN+ WE LA +S H ETG PLPKE ++L AKNFQ  
 Sbjct: 485 VSDVAGINGVPWDAVELPSQFMENWCWEEELAFISGHYETGFLPKKELTQLLAKKNFQ 544

Query: 541 RGMFLVVRQMEFALFDMMIYSEDDEGRKKNWQQVLDVSRKEVAVVRPPEYNRFANSFGHIF 600  
 MF++RQ+EF +FD ++ D + L SV+ +VAV++ ++ R +SF HIF  
 Sbjct: 545 AAMFILRQLEFGIFDFRLHHTFDAEKTNQILDTLKSVKSQVAVIKGVDMWARAPHFSHF 604

Query: 601 XXXXXXXXXXXXWAEVLSADAYAAFEESDDV-AATGKRFWQEILAVGGSRSAAESFKAFR 659  
 WAEVLSADAY+ FEE TKG F EIL GGS E FK FR  
 Sbjct: 605 AGGYAAGYYSYLWAEVLSADAYSRFEEEGIFNPITGKSFLDEILTRGSGSEPMELFKRFR 664

Query: 660 GREPSIDALLRHSGFDN 676  
 GREP +DALLRH G N  
 Sbjct: 665 GREPQLDALLRHKGIMN 681

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 493>:

g129.seq  
 1 ATGCTTTCAC CTCCTCGGCG TAAACGGCG GCACATCAAT CAAGCCGCTC  
 51 TTCATTGCGG TGCGGAAAAA ATGCGGCGTG TTGCCGTGAT CAAAATCAAT  
 101 ATCGTGCAGC ATCCAGCCCA AATCGCGGTT TGCTCGCTT TCCGATAACG  
 151 CCGACGGCGG CAGCGGTTCA CCCTTATCCG CGCTTTCGCC ATTTGCCCTT  
 201 TCAGGCTGCG GGCATAGGGG CGGAACAGGC GCGGTCGAA TCCTGTTTCA  
 251 TCCGACAAA CGCGTTGGCA GTCGAAAAAT CCGGCCGGCC GTGTCAAATA  
 301 ATGCGTTACT TTGGCCGGGT CTGTCTCTTT GTAAGCGGCG GTCTTTTTTT  
 351 GCGCGCCATC CGCATCTGTT TGGGCGCATG GCAAACGGCG GCTGCCGTAC  
 401 AATCAAAATG TTTGGCGATT TCATGCAGAC AGGCATCCG ATGCCGCCG  
 451 ACATATCGAG CCGGTTTTTG CCTATCCGAT TTGGCGGCAT TTAGCCCGGT  
 501 AACTTGA

This corresponds to the amino acid sequence <SEQ ID 494; ORF 129.ng>:

g129.pep  
 1 MLSPPRRKTA AHQSSRLSFA CGKNAACCRD QNQYRAASSP NRGLPRFPIT  
 51 PTAAAVHPYP RFRHLPFQAA GIGAEQAAVE SCFIRTNALA VGKSGRPQI  
 101 MRYFGRVLSF VSGGLFLRAI RICLGAWQTA AAVQSKCLAI SCRQASGCRP  
 151 TYRAGFCLSD LAAFRPVT\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 495>:

m129.seq (partial)  
 1 ..TATCTGCGCT TTTACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA  
 51 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG  
 101 GAAATTCGG CCGGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG  
 151 TTCTTTGTAA GTGGTGGTCT TTTTTCGCG GTTATCCCCA TCTGTTTGAG  
 201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTTCA  
 251 GCAGATAGGC ATCCGGGTGT TGCCCAACAT ATTAGCCCG TTTTTCCTA  
 301 TCCGATTTGA CGGCATTTAG ACCGGTAACT TGA

This corresponds to the amino acid sequence <SEQ ID 496; ORF 129>:

m129.pep (partial)  
 1 ..YLRFHLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGRLC QIMRYFGRVL  
 51 FVSGGLFLR VIPICLSAXQ MVAAVQSKCL AISCRXASGC CPTYXAGFCL  
 101 SDLTAFRPVT \*

Computer analysis of this amino acid sequence gave the following results:

375

```

251 AATGCCACGC GCGGACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGACT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCTCA
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGCAG
401 ACCTGACCGA TCAGGAACTC AAACGGGCGA TTACCTACAT GGCGAATAAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGAGCT CCGGCAGAAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651 CGGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAAGA AACCTTGCAC AAACATGCCC TTGAAGGCTT TAACGCGATG
751 CCGGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801 TGTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 500; ORF 130.ng>:

```

g130.pep
1  MKQLRDNKAQ GSALFTLVSG IVIVIAVLYF LIKLAGSGSF GDVDATTEAA
51  TQTRIQPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101 NGDWAPRIAQ GFDTLFQHAL NGFNAMPAKG GAADLTQDEL KRAITYMANK
151 SGGSFNPDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAPAVGVGD
201 KKVFEATCQV CHGGSIPGIP GIGKDDWAP RIKKGKETLH KHALEGFNAM
251 PAKGGNAGLS DDEVKAAVDY MANQSGAKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 501>:

```

m130.seq (partial)
1  ..GGCGAACAGA TTTTCGGCAA AATCTGTATC CAATGCCACG CCGCGGACAG
51  CAATGTGCCG AACGCTCCGA AACTGGAACA CAACGGCGAT TrGGCACCGC
101 GTATCGGCAA GGCTTCGATA CCTTGTTCCTA ACACGCGCTG AACGGCTTTA
151 ACGCCATGCC TGCAAAAGGC GGTGCGGCAG ACCTGACCGA TCAGGAACTT
201 AAACGGGCGA TTACTTACAT GGCGAACAAA AGCGGCGGTT CTTTCCCGAA
251 TCCTGATGAG GCTGCGCCTG CCGACAATGC CCCTTCAGGA ACAGCTTCTG
301 CTCCTGCCGA TAGTGACGCT CCGGCAGAAG CGAAGGCAGA AGACAAGGGT
351 GCGGCACCCC TGCGGTCGGC GTTGACGTA AAAAAAGTCTT CGAAGCAACC
401 TGTGAGGTGT GCCACGGCGG TTCGATTCCC GGTATTCCCG GCATAGGCAA
451 AAAAGACGAT TGGGCACCGC GTATCAAAAA AGGCAAAAGAA ACCTTGCACA
501 AACACGCCCT TGAAGGCTTT AACGCGATGC CTGCCAAAG CGGCAATGCA
551 GGTTTGAGCG ATGACGAAGT CAAAGCGGCT GTTGACTATA TGGCAAACCA
601 ATCCGGTGCA AAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 502; ORF 130>:

```

m130.pep (partial)
1  ..GEQIFGKICI QCHAADSNVP NAPKLEHNGD XAPRIQGFDT LFQHALNGFN
51  AMPAKGGAAD LTDQELKRAI TYMANKSGGS FPNPDEAAPA DNAASGTASA
101 PADSAAPAEA KAEDKGAAPA VGVDGKKVFE ATCQVCHGGS IPGIPIGKK
151 DDWAPRIKKG KETLHKHALE GFNAMPKXG NAGLSDDEVK AAVDYMANKS
201 GAKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 130 shows 98.1% identity over a 206 aa overlap with a predicted ORF (ORF 130.ng) from *N. gonorrhoeae*:

```

m130/g130
10 20 30
m130.pep GEQIFGKICIQCHAADSNVPNAPKLEHNGD
|||||
g130 DATTEAATQTRIQPVGQ LTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
50 60 70 80 90 100

40 50 60 70 80 89
m130.pep XAPRI-QGFDTLFQHALNGFNAMPAKGGAADLTQELKRAITYMANKSGGSFPNPDEAAP
|||||
g130 WAPRIAQGFDTLFQHALNGFNAMPAKGGAADLTQELKRAITYMANKSGGSFPNPDEAAP
110 120 130 140 150 160

```

```

      |||
a130  KKDDWAPRIKKGKETLHKHALEGFNAMPKGGNAGLSDDEVKAAVDYMANQSGAKFY
      230      240      250      260      270      280

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 505>:

```

g132.seq
  1  ATGGAAGCCT TCAAAACCTT AATTGGATT ATTAATATTA TTTCCGCTTT
 51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101  GCGCGACCTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151  GCCGGCAACG CCAACTTcct CAgccGCTCG AccGccGTTG CAGCAACAtt
201  tttctttTGca acctgcAtgg gctatggTgt atattcacac CCACACGACA
251  AAACACGGTT TGGACTtcag caacataCGA CAGACTCAGC AagcACCCAA
301  ACCcgtAAGC AATACCGAAC CTTCTGCCCC TGTTCTCTCAG CAGCAGAAAT
351  AACagtTTTT CAAATgccga catGgtga

```

This corresponds to the amino acid sequence <SEQ ID 506; ORF 132.ng>:

```

g132.pep
  1  MEAFKTLIWI INIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
 51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QHTTDSASTQ
101  TRKQYRTFCP CSSAAEITVF QMPTW*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 507>:

```

m132.seq (partial)
  1  ATGGAACCCT TCAAAACCTT AATTGGATT GTTAATTTAA TTTCCGCTTT
 51  GGCCGTCCTC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101  GCGCGACTTT CGGA...

```

This corresponds to the amino acid sequence <SEQ ID 508; ORF 132>:

```

m132.pep (partial)
  1  MEPFKTLIWI VNLISALAVF VLVLLQHGKG ADAGATFG...

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 132 shows 89.5% identity over a 38 aa overlap with a predicted ORF (ORF 132.ng) from *N. gonorrhoeae*:

```

m132/g132
      10      20      30
m132.pep  MEPFKTLIWI VNLISALAVF VLVLLQHGKG ADAGATFG
      || |||||:|:|||||:|||||
g132      MEAFKTLIWI INIISALAVI VLVLLQHGKG ADAGATFGSGSGSAQGVFGSAGNANFLSRS
      10      20      30      40      50      60

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 509>:

```

a132.seq
  1  ATGGAAGCCT TCAAAACCTT AATTGGATT GTTAATATAA TTTCCGCTTT
 51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101  GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151  GCCGGCAACG CTAACCTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
201  TTTCTTTGCA ACCTGCATGg GCTATGGTGT ATATTCACAC CCACACGACA
251  AAACACGGTT TGGACTTCAG CAACGTACAA CAACTCAGC AAGCACCCAA
301  ACCCGTAAGC AATACCGAAC CTTCTGCCCC TGTTCTCTCAG CAGCAGAAAT
351  AACAGTTTTT CAAATGCCGA CATGGTGA

```

This corresponds to the amino acid sequence <SEQ ID 510; ORF 132.a>:

```

a132.pep
  1  MEAFKTLIWI VNIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
 51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QRTTNSASTQ
101  TRKQYRTFCP CSSAAEITVF QMPTW*

```

m132/a132 92.1% identity in 38 aa overlap

```

201 TTCCGTGGCA TCAAGTGTGA TGCAGTTCGA TTACAAAGAC CACACCGTCA
251 ACCTCTTGGA CACGCCGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301 GTTTTAACCG CCGTGGACAG CGCATTAAATG STCATCGACG CGGCAAAAGG
351 CGTGGGAAGCG CAAACCATCA AGCTCTTAAA CGTCTGCCGC CTGCGCGATA
401 CACCGATTGT TACGTTTATG AACAAATACG ACCGCGAAGT GCGCGATTCC
451 CTGGAACCTT TGGACGAAAGT GGAAAACATT TTAAAAATCC GCTGCGCGCC
501 CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAGGGG GGTGTACCACA
551 TCCTGAACGA TGAAATTTAT CTCTTTGAAAG CTGGCGGCGA ACGCCTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCTGAAT TGGAAACAACG
651 CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGGCTTCCAA CGAGTTTAAT CTCGACGAAT TCCTCGCCCG CGAACTCACG
751 CCCGTATTCT TCGGCTCTGC GATTAACAAC TTGGGTATTC AGGAAATCCT
801 CAATTTCATT ATTGACTGGG CGCCCGCGCC GAAACCGCGC GACGCGACCG
851 TACGTATGGT CGAGCCGGAC GAGCCGAAGT TTTCCGGATT TATCTTCAAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CTTCTTGCG
951 CGTCTGTCCG GGCAAAATCG AGCGCGGCAT GAAGATGAAA CACCTCGCTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TTACCTTCAT GTCGCACGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCGGCG GACATTATCG GCATCCGAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACCTG
1151 CGTTCACCGG CATCCCATTC TTCGCACCCG AACTGTTCCG CAGCGTACGC
1201 ATCAAAAACC CGCTGAAAAT CAAACAACCTG CAAAAAGGCT TGCAACAGCT
1251 CGGCGAAGAA GGCGCGGTGC AGGTGTTCAA ACCGATGAGC GCGCGCGATT
1301 TGATTTTGGG CGCGGTCGGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351 CTCGCCAACG AATACGGCGT AGAAGCCGTG TTCGACAGCG CATCCATCTG
1401 GTCGCGCGCG TGGGTATCGT GCGACGACAA GAAAAAATG GCTGATTTG
1451 AAAAAAGCAA CGCGGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501 TACCTCGCCC CCAACCGCGT GAATTTGGGA CTCACGCAAG AACGTTGGCC
1551 GGACATCGTG TTCCACGAAA CACGCGAACA TTCGGTCAAA CTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 514; ORF 134>:

```

m134.pep
1 MSQEILDQVR RRRTFAIISH PDAGKTLTE KLLLFSGAIQ SAGTVKGKKT
51 GKFATSDWME IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151 LELLDEVENI LKIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGVAVV FDSASIWSAR WVSCDDKKKL AEFEKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 134 shows 98.7% identity over a 531 aa overlap with a predicted ORF (ORF 134.ng)

from *N. gonorrhoeae*:

```

m134/g134
      10      20      30      40      50      60
m134.pep MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
|||||
g134      MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD
      10      20      30      40      50      60

      70      80      90     100     110     120
m134.pep IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDLSALMVIDAAKGVEA
|||||
g134      IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDLSALMVIDAAKGVEA
      70      80      90     100     110     120

      130     140     150     160     170     180

```

381

```

951 CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TAACCTTCAT GTCCACAGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCGGCG GACATTATCG GTATCCCAAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACCTGA
1151 CGTTTACCGG CATCCCATTC TTGCGGCGCG AACTGTTCCG CAGCGTTCGC
1201 ATCAAAAACC CGCTGAAAAT CAAGCAACTG CAAAAAGGTT TGCAACAGCT
1251 TGGCGAAGAA GGTGCGGTGC AGGTGTTCAA ACCAATGAGC GGCGCGGATT
1301 TGATTTTGGG CGCGGTCGGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351 CTTGCCAAGC AATACGGCGT GGAAGCCGTG TTCGACAACG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAAGT GCGGAATTTG
1451 AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCGGGCGG CAACCTCGCC
1501 TACCTCGCCC CTAACCGCGT GAATCTGGGA CTCACGCAAG AACGCTGGCC
1551 GGACATCGTG TTCCACGAAA CGCGCGAGCA TTCGGTCAAA CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 516; ORF 134.a>:

```

a134.pep
1 MSQEILDQVR RRRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
51 GKFDATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAGVEA QTIKLLNVCR LRNTPIVTFM NKYDREVRDS
151 LELLDEVENI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGID NPELEQRFPL EIQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IEWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLTFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGEAV FDNASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

m134/a134 98.9% identity in 531 aa overlap

```

          10      20      30      40      50      60
m134.pep MSQEILDQVRRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
          |||
a134      MSQEILDQVRRRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD
          10      20      30      40      50      60

          70      80      90      100     110     120
m134.pep IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDLSALMVIDAAGVEA
          |||
a134      IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDLSALMVIDAAGVEA
          70      80      90      100     110     120

          130     140     150     160     170     180
m134.pep QTIKLLNVCRRLDTPIVTFMKNKYDREVRDSLELLDEVENILKIRCAPVTWPIGMGKNFKG
          |||
a134      QTIKLLNVCRRLDTPIVTFMKNKYDREVRDSLELLDEVENILQIRCAPVTWPIGMGKNFKG
          130     140     150     160     170     180

          190     200     210     220     230     240
m134.pep VYHILNDEIYLFAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
          |||
a134      VYHILNDEIYLFAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
          190     200     210     220     230     240

          250     260     270     280     290     300
m134.pep LDEFLAGELTPVFFGSAINNFGIQLNSLIEWAPAPKPRDATVRMVEPDPEPKFSGFIFK
          |||
a134      LDEFLAGELTPVFFGSAINNFGIQLNSLIEWAPAPKPRDATVRMVEPDPEPKFSGFIFK
          250     260     270     280     290     300

          310     320     330     340     350     360
m134.pep IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
          |||
a134      IQANMDPKHRDRIAFLRVCSGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
          310     320     330     340     350     360

```



```

501 GACCGACATA GACGGTCTTT ACACGGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCGGC AAACGGCAGC GGCGGTATGC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CTGCCGAACA TCAGGCGGAC
751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATCCGAAA GCCGGGGCAG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTGTCCGAA CAGGGGAAAA GCCTGCTGAT GTCGGGCATT
901 GCCGGAATCG AAGGGCATT TTCCCGTATG GACACCGTAA CCGTGTACAG
951 CAAGGCAACC AAACAGCCCC TGGGCAAAGG GCGCGTCTTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATCGCGTA AGGCGAAAGG CGTGTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTTAG

```

This corresponds to the amino acid sequence <SEQ ID 520; ORF 135>:

```

m135.pep
1  MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TCQLAALHHA GHVLVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAPV IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHNHEIEM
201 AGGSGSANGT GGMLTKIKAA TIAESGVPV YICSSLKPDA LAEAAEHQAD
251 GSFFVPRAGK LRTQKQWLAF YSESRGSVYV DEGAEHALSE QGKSLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAEDLL KSRKAKGVFI
351 HRDDWISITP EIRLLLTEF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 135 shows 97.6% identity over a 294 aa overlap with a predicted ORF (ORF 135.ng)

from *N. gonorrhoeae*:

```

m135/g135
      10      20      30      40      50      60
m135.pep  MKYKRIVFKVGTSSITHSDGSLSRGKIQTI TCQLAALHHAGHELVLVLVSSGAVAAGFGALG
          |||
g135      MKYKRIVFKVGTSSITRSDGSLSRGKIQTI TRQLAALHHAGHELVLVLVSSGAVAAGFGALG
          |||
      10      20      30      40      50      60

      70      80      90      100     110     120
m135.pep  FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLSRADFADKRRYQNAGGAL
          |||
g135      FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQILLSRADFADKRRYQNAGGAL
          |||
      70      80      90      100     110     120

      130     140     150     160     170     180
m135.pep  SVLLQRRAPVPIINENDTVSVEELKIGDNDT LSAQVAAMIQADLLVLLTDIDGLYTGNPNS
          |||
g135      SVLLQRRAPVPIINENDTVSVEELKIGDNDT LSAQVAAMIQADLLVLLTDIDGLYTGNPNS
          |||
      130     140     150     160     170     180

      190     200     210     220     230     240
m135.pep  NPDAVRLDKIEHNHEIEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPDA
          |||
g135      NPDAVRLDKIEHNHEIEMAGGSGSANGTGGMLTKIKAATIAESGVPVYICSSLKPD
          |||
      190     200     210     220     230     240

      250     260     270     280     290     300
m135.pep  LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRGSVYVDEGAEHALSEQGKSLMSGI
          |||
g135      LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESGSVYVDESAEHALSEQGKACX
          |||
      250     260     270     280     290

      310     320     330     340     350     360

```

385

	250	260	270	280	290	300
m135.pep	LAEAAEHQADGSFFVPRAKGLRTQKQWLAFYSES	SRG	SVYVDEGA	EAH	LS	EQGKSLLMSGI
	::					
a135	LAEAADNQADGSFFVPRAKGLRTQKQWLAFYSES	RG	GVYVDEGA	EAH	LS	EQGKSLLMSGI
	250	260	270	280	290	300
	310	320	330	340	350	360
m135.pep	AGIEGHFSRMDTVTVYSKATKQPLGKGRVLF	GSAA	EDLLKSR	KAGVFI	HRDD	WISITP
a135	AGIEGHFSRMDTVTVYSKATKQPLGKGRVLF	GSAA	EDLLKSR	KAGVFI	HRDD	WISITP
	310	320	330	340	350	360
	370					
m135.pep	EIRLLLTEFX					
	(     )					
a135	EIRLLLTEFX					
	370					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 523>:

```

g136.seq
1      ATGGAATACC GGTTCAGAC AGCATTTTFA CGTTTGGTTC AGatgaAAAC
51     AAACGCTca aTTCTtacg caACACGGCT TGTATTTCCT GcgcCTGCCG
101    CACGGACAGG GATCGTTTCT GCCGgtTTTT TCCCCTTCCC TGCGGACGGT
151    TTGCGGTTTG TTGATGACCG CCTGCCAGTA GCGGTAGATG Tctgccagcg
201    cgTAAGGCag tTCGGAGcga agtctccgta gctcgcttcc ggTGAATTGC
251    AGcgcgataa cgcgcttTC TCTTCGTCG taaatgccgc ccaactgcat
301    cacgGGGTAA AACAGCTCTT CAAACGCTTC ATCATCGGCG GCTTCAAACC
351    AATCGGTTCG CACAATGTCC AAACCGCTAA GATAGGCGTT GCACCAAGTG
401    TAAAAATCGC TGCCGCGCTC CGCGTCGTCG TAGAGCCACA AATCAAGGTG
451    CTTTTTATCC GACATCGCGG GCGTGTGTTT CATGCGCAAT GCCAAAACCA
501    GCCGTTTCGAT TTCGGAACGT TCGGCGGCGG TAAATTGCGA TTCGTGCCCC
551    AACACTTCGC GCAGCCAGTC GAGGCGTGCC AATTGTGCTG GCCCGTCAA
601    CAGCGCCGTC ATAAACCTT GAACCTCGTC GCAACGCATC GTGTGTGCTT
651    GTTCGCTTTT GGCATCCAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 524; ORF 136.ng>:

```
g136.pep
1  MEIRFQTAFL RLVQMKTNAS ILTATRLVFP AAAARTGIVP AGFFPPPADG
51  LRFVDDRLPV AVDVQQRVRQ FGRKFRQLAF GELQADNAV F LFVNVAAHCH
101 HGVKQLPKRF IIGGFKPIGR HNVOTVKIGV APSVKIAAAL AVVVPEQTIG
151 LFIHRHGGCF HRHCQNQVAF GTFTGGGKLR FVAQHFGQP V ERCQFVRPAQ
201 QRRHKTNLNV ATHRQALPFA FGQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 525>:

```
m136.seq
1   ATGGAAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC
51  CGCTGCGCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCTGT
101 CGGACGGTTT GCGGTTTTGT GATGACTGCC TGCCAGTAGC GGCTAGATATC
151 CGCCAATGCA TAAGGCAACT CGSATTCCAG TTCGCCCAGC TCGCCTTCTG
201 TGAATTGCAG ACGGATAGCG CCGTTTTCTT CTTCGTCGTA AATACCGCCC
251 AATGCCCATGA TGGGATAAAA CAACTCTTCA AACGCTTCAT CATCGACGGC
301 TCAAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATATGC
351 ACCATGTGTA AAAATCGCTG CCGCCGCTCT CGTTTTCATA CAGCCACAAA
401 TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTTCCA TCGCGATGAT
451 CAAACACCAG CGTTCGATTG CGGAACGTTT GCGCGCGGTA AATTGCGATT
501 CGTCGCCCAA CACTTCGGGC AGCCAGTCEA GCGGTGTCAA TTTGTCCGGC
551 CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT
601 GTTGCCTTGT TCGCTTTTGG CATCCAACAA TTGCTCAAC CGCGTTTTGG
651 ATGCTTCGGT AAATTTTTCG GAATCCATCA TTTTCTTTT CAAATGGGTT
701 TTGCGCCCTA TTATCGCCGC AATGCGGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 526; ORF 136>:

387

```

1 METNASILTA TRLVFSAAAA RTGIVPACFF AFPADGLRLV DDRLPVAVDI
51 RQCIRQLGFQ FRQLAFCELO TDSAVVLFV NTAQCHDGIK QLFKRFIIDG
101 FKPIGRHNIQ TVKISIAPCV KIAA VFVEI QPQIGQFFIR HRGGCFHRHC
151 QNQPFDFGTG GGGKLRFVAQ HFGQPVERCQ FVRPAQRRH KTLNLVATHR
201 VALFAFGIQQ FAQPPFGCFG KFSGIHHFPF PMGFAPYYSG LNLNQDKATK
251 PQTVQIVRQG EATPYWFKFN PLYRRNAV*

```

m136/a136 98.3% identity in 238 aa overlap

```

10 20 30 40 50 60
m136.pep METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRFVDDCLPVAVDIRQCIRQLGFQ
|||||
a136 METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRLVDDRLPVAVDIRQCIRQLGFQ
10 20 30 40 50 60

70 80 90 100 110 120
m136.pep FRQLAFCELO TDSAVVLFVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV
|||||
a136 FRQLAFCELO TDSAVVLFVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV
70 80 90 100 110 120

130 140 150 160 170 180
m136.pep KIAA VFVEIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ
|||||
a136 KIAA VFVEIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLRFVAQHFGQPVERCQ
130 140 150 160 170 180

190 200 210 220 230 240
m136.pep FVRPAQRRHKTNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPFPMGFAPYYRR
|||||
a136 FVRPAQRRHKTNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPFPMGFAPYYSG
190 200 210 220 230 240

m136.pep NAVX
a136 LNLNQDKATKPQTVQIVRQGEATPYWFKFNPLYRRNAVX
250 260 270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 529>:

```

g137.seq
1 ATGATTATCC ATCACcaATT CGATCCCGTC CTCATCAGTA TCGGCCCGCT
51 TGCCGTCCGC TGGTATGCCT TAAGCTACAT CCTCGGATT ATTCTTTTAA
101 CCTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCTG TTTTACCAAA
151 GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGCGG TGATTTTGGG
201 CGGACGCTTG GGCTATGTCC TGTTTTACAA ATTTCTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTGG GTGTAGTTAT TGCCATATGG TTGTTAGCC GCAAGCACGG
351 CATCGGCTTC CTCAAAGTGA TGGACACGGT CGCGCCGCTC GTTCCGCTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTTTA TCAACGGCGA ACTTTGGGGA
451 CGCATTACCG ACATTAACGC ATTTTGGGCA ATGGGCTTCC CGCAAGCGCA
501 TTACGAAGAT GCCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCCCTTGAAG GCATCTGCCT GTTCGCGGTC GTTTGGCTGT TTTCCAAAAA
651 ACCGCGCCCG ACCGGGCGA CTGCGCGGCT TTTTCTCGGC GGCTACGGCG
701 TGTTCCGCTT TATTGCCGAA TTTGCGGCGC AACCCGACGA CTATCTCGGG
751 TGCTGACCT TGGGGCTGTC GATGGGCAAA TGGTTGAGCG TCCCGATGAT
801 TGTTTTGGGT ATCGTCGGCT TTGTCGGGTT CGGCATGAAA AAACAGCACT
851 GA

```

This corresponds to the amino acid sequence &lt;SEQ ID 530; ORF 137.ng&gt;:

g137.pep

```

g137      |||||
          AEWLQQYGM LPRHPSQLYQFALEGICLFAVVWLFSSKKPRPTGQTAAFLGGYGVFRFIAE
          190      200      210      220      230      240

          250      260      270      280
m137.pep  FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX
          |||||
g137      FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX
          250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 533>:

```

a137.seq
1  ATGATTACCC ATCCCAATT CGACCCCGTC CTTATCAGTA TCGGCCCGCT
51  TGCCGTCCGC TGGTATGCCC TAAGCTACAT CCTCGGATTT ATTCTTTT
101 CCTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCGGT TTTTACCAA
151 GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TAATTTGGG
201 CGGGCGTTTG GGTACGTCCT TGTTTTACAA GTTTTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTGG GTGTAGTTAT TGCCATATGG TTGTCGGTC GCAAACACGG
351 CATCGGCTTC CTAAACTGA TGGACACGGT CGCACCGCTC GTTCCACTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTCA TCAACGGCGA ACTTTGGGGA
451 CGCGTTACCG ACATCAACGC ATTTTGGGCA ATGGGCTTCC CGCAGGCGCG
501 TTACGAAGAC CTCGAAGCCG CCGCGCACA TCCGCTTTGG GCAGAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCACTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TCTCTAAAAA
651 ACAGCGGCGC ACCGGACAAG TCGCCTCACT CTTCTCGGC GGCTACGGCA
701 TATTCGCTT CATTGCCGAA TTGACACGCC AACCCGACGA CTATCTCGGG
751 CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCAGATGAT
801 TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AACAGCACT
851 GA

```

This corresponds to the amino acid sequence <SEQ ID 534; ORF 137.a>:

```

a137.pep
1  MITHPQFDPV LISIGPLAVR WYALSYILGF ILETFGLRRR IAQGLSVFTK
51  ESLDDFLTWG ILGVILGGRL GYVLFYKFS DYL AHP LDI FK V WEGGMSFHG
101 GFLGVVIAIW LFGRKHGIGF LKLM DTVAPL VPLGLASGR I GN FINGELWG
151 RVT DINA FWA MGFPQARYED LEAAAHNPLW AEWLQQYGM L PRHPSQLYQF
201 ALEGICLFAV VWLFSSKKQRP TGQVASLFLG GYGIFRFIAE FARQPDDYL G
251 LTLGLSMGQ WLSVPMIVLG IVGFVRFGMK KQH*

```

m137/a137 98.2% identity in 283 aa overlap

```

m137.pep      10      20      30      40      50      60
MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
|||||
a137          10      20      30      40      50      60
MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG

          70      80      90      100     110     120
m137.pep  ILGVILGGRLGYVLFYKFS DYLAHPLDIFKVWEGGMSFHGGFLGVVIAIRLFGRKHGIGF
|||||
a137      ILGVILGGRLGYVLFYKFS DYLAHPLDIFKVWEGGMSFHGGFLGVVIAIWLFGRKHGIGF
          70      80      90      100     110     120

          130     140     150     160     170     180
m137.pep  LKLM DTVAPL VPLGLASGR I GN FINGELWGRVTDINAFWAMGFPQARYEDAEAAAHNPLW
|||||
a137      LKLM DTVAPL VPLGLASGR I GN FINGELWGRVTDINAFWAMGFPQARYEDLEAAAHNPLW
          130     140     150     160     170     180

          190     200     210     220     230     240
m137.pep  AEWLQQYGM LPRHPSQLYQFALEGICLFTVIWLFSSKKQ RSTGQVASLFLGGYGIFRFIAE
|||||
a137      AEWLQQYGM LPRHPSQLYQFALEGICLFAVVWLFSSKKQ RPTGQVASLFLGGYGIFRFIAE

```

```

51 RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKTEAMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
151 DIGQVGTVES IDTGLVKGLI ERGCI PVVAP VGVGEKGEAF NINADLVAGK
201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
251 KIASAVEAAV NGVKATHIID GRPLNALLE IFTDAGIGSM ILGGGEDA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 138 shows 98.0% identity over a 298 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae*:

m138/g138

	10	20	30	40	50	60
m138.pep	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
g138	MEFENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m138.pep	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKTEAMDIVEMVLGGHVNKEIVSMINTY					
g138	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKETMDIVEMVLGGHVNKEIVSMINTY					
	70	80	90	100	110	120
	130	140	150	160	170	180
m138.pep	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCI PVVAP					
g138	GGHAVGVSGRDDHFIKAKKLLVDTPQNSVDIGQVGTVESIDTGLVKGLIERGCI PVVAP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m138.pep	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
g138	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDGLIA					
	190	200	210	220	230	240
	250	260	270	280	290	299
m138.pep	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRPLNALLEIFTDAGIGSMILGGGEDAX					
g138	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRPLNALLEIFTDAGIGSMILGRGEDAX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 539>:

a138.seq

```

1 ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
51 CGAAGCGCTG CCTTACATCC GCCGGTTTTC CGGTCGGTC GCCGTCATCA
101 AATACGCGCG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGA AAAAGTC GGCAAAAAGG
251 GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAGAGGC GATGGATATT
301 GTCGAAATGG TGTGGGCGG GCATGTCAAT AAAGAAATCG TGTCGATGAT
351 TAACACATAT GCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
401 TCATTAAGGC GAAGAACTT TTGATCGATA CGCCGAACA GAATGGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT AGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAACTC TTGATGATGA CGAATATCGC
651 CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAATC ACGCCGAAAC
701 GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCCGCGTC AACGGCGTGA AAGCCACGCA
801 TATCATCGAC GGCAGGGTGC CCAACGCGCT TTTGCTGGAA ATCTTACCG
851 ATGCCGGTAT CGGTTCGATG ATTTTGGGCG GTGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 544; ORF 138>:

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 138 shows 92.2% identity over a 179 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae*:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 545>:

This corresponds to the amino acid sequence <SEQ ID 546; ORF 139.a>:

a139.pep  
1 MRTTPTEPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS

395

251 DSISLFAGIR HDVGDIGYLK GLFSYGRYKN SISRSTGADE YAEGSVNGTL  
 301 MQLGALGGVN VPFAATGDLT VEGGLRHDLL KQDAFAEKGS ALGWSGNSLT  
 351 EGTLVGLAGL KLSQPLSDKA VLSATAGVER DLNGRDYAVT GGFTGAAAAT  
 401 GKTGARNMPH TRRVAGLGVD VEFNGWNGL ARYSYTGSKQ YGNHSGQIGV  
 451 GYRF\*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 549>:

m140.seq  
 1 ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG  
 51 TGTTCCTTC CTGAGTGGCG CCAAAATCGG GCAGGATTAT TCTTTCTTCA  
 101 CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCCTCGA CAGCGTCGAA  
 151 AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG  
 201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC  
 251 TGAACACGCG CGTAGAACAG GCGCGCAGCA ATCTGGAAAA CCTGATGGTC  
 301 GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC  
 351 GGCAGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCTAC GGCACAACTT  
 401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCGA CGGTGTACGC  
 451 ATCTTCAACA GTCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA  
 501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC  
 551 ACAACGGCAC GGGTCTGCGC GTCATCGCGC AAACCCAAACA GGACGGTGGA  
 601 ACGTGGGAAC AGGGCGGTGT TGAAGCAAA ATGCGCGGCA GTACCCAAAC  
 651 CGTCGGCATT GCCCGGAAAA CCGCGGAAAA TACGACAGCA GCCGCCACAC  
 701 TGGGCATGGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC  
 751 GACAGCATTG GTCTGTTTGC AGGCATACGG CACGATCGCG GCGATATCGG  
 801 CTATCTCAA GGCCTGTTCT CTACGACG CTACAAAAAC AGCATCAGCC  
 851 GCAGCACCGG TGCGGACGAA CATGCGGAAG GCAGCGTCAA CCGCACGCTG  
 901 ATGCAGCTGG GCGCACTGGG CGGTGTCAAC GTTCCGTTT CCGCAACGGG  
 951 AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG  
 1001 CATTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCTCACT  
 1051 GAAGGCACGC TGGTCCGACT CGCGGGTCTG AAGCTGTGCG AACCTTGAG  
 1101 CGATAAGGCC GTCCTGTTTG CAACGGCGGG CGTGGAAACG GACCTGAACG  
 1151 GACGCGACTA CACGGTAACG GCGCGCTTTA CCGCGCGGAC TGCAAGCAACC  
 1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGTCTGG TTGCCGGCCT  
 1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA  
 1301 GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA  
 1351 GGCTACCGGT TCTGA

This corresponds to the amino acid sequence <SEQ ID 550; ORF 140>:

m140.pep  
 1 MSARGKGAGY LNSTGRRVPF LSAAKIGQDY SFFTNIETDG GLLASLDSVE  
 51 KTAGSEGDTL SYVRRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV  
 101 ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR  
 151 IFNSLAATVY ADSTAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQDGG  
 201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT  
 251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL  
 301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT  
 351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT  
 401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV  
 451 GYRF\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 140 shows 94.5% identity over a 454 aa overlap with a predicted ORF (ORF 140.ng) from *N. gonorrhoeae*:

m140/g140

	10	20	30	40	50	60
m140.pep	MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL					
g140	MSARGKGAGYLNSTGRHVPFLSAAKIGQDYSFFKNIKTDGGLLASLDSVEKTAGSEGDTP					
	10	20	30	40	50	60
	70	80	90	100	110	120

```

951 AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG
1001 CATTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCATCACT
1051 GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAGCTGTCGC AACCCTTGAG
1101 CGATAAGCC GTCCTGTTG CAACGGCGGG CGTGAACGC GACCTGAACG
1151 GACGCGACTA CACGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGCCTGG TTGCCGGTCT
1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301 GCTACGCCGG TTCCAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
1351 GGCTACCGGT TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 552; ORF 140.a>:

```

a140.pep
1 MSAGGKGAGY LNRTGQRPVF LSAAKIGRDY SFFTNIETDG GLLASLDSVE
51 KTAGSEGDTL SYVVRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
101 ELDAESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR
151 IFNNLAATVY ADSTAAHADM QGRRLLKAVSD GLDHNATGLR VIAQTQDGG
201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGHST WSENSANAKT
251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSIT
351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV
451 GYRF*

```

m140/a140 98.2% identity in 454 aa overlap

	10	20	30	40	50	60
m140.pep	MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGD					
a140	MSAGGKGAGYLNRTGQRPVFLSAAKIGRDYSFFTNIETDGGLLASLDSVEKTAGSEGD					
	10	20	30	40	50	60
	70	80	90	100	110	120
m140.pep	SYVVRGNAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDAESSATPETVETAAAD					
a140	SYVVRGNAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDAESSATPETVETAAAD					
	70	80	90	100	110	120
	130	140	150	160	170	180
m140.pep	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLLKAVSD					
a140	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNNLAATVYADSTAAHADMQGRRLLKAVSD					
	130	140	150	160	170	180
	190	200	210	220	230	240
m140.pep	GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGRST					
a140	GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m140.pep	WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNISISRSTGADEHAEGSVNGTL					
a140	WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNISISRSTGADEHAEGSVNGTL					
	250	260	270	280	290	300
	310	320	330	340	350	360
m140.pep	MQLGALGGVNVFPAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSITEGTLVGLAGL					
a140	MQLGALGGVNVFPAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSITEGTLVGLAGL					
	310	320	330	340	350	360
	370	380	390	400	410	420
m140.pep	KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD					
a140	KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD					



399

m141.seq

```

1  ATGAGCTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTTGGTCT GAATGCCGAC AACATTGAGC
101 CTTACGGTCA TTACAAGGCG AAAATCAATC CTGCCGAAGC GTTCAAAGTG
151 CCGCAAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
201 GGCGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCGTTGC
251 GCCACATCGG CAAAGATGCC GTGATTGCC TGGCGAACC TTCTCTGGGG
301 CCGGTGTTTC GCGTGAAAGG CGGCGCGGCA GGCGCGGCT ATGCCCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
401 TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAA CGCGTGCTGT GGCGGCGCGT
501 GGTTCGATAT AACGACCGCC AGTTGCGCAA CATCATC3AC GGCATGGGTA
551 AACCCTTTGA CGGCGTGATG CGTCTGACG GTTTCGATAT TACCCTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTTTG GGCAACATCC TTGTCCGCTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
751 GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCGCGCTT
801 CGTACACGGC GGCCCGTTTC CCAACATCGC CCACGGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCGCT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CCTTGCCGGT TTGAAACCTG ATGCGGCTGT TGTCTGCGC ACTGTCCGCG
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAT
1051 TTAGACGCTT TGGAAAAAGG TTTGCCAAC CTGCTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTCGAC TGCCCGTCGT CGTTGCGCTC AACCCTTCG
1151 TGTCCGACGC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CTGCGCGTTG AAGTTTCCCT GACCGAAGTG TGGGCGAAAG GTGGTGCGGG
1251 CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
1301 ATAACCTTCG TTTGCGCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCAAAAAAGT GTACGGCGCG GAAGATGTTG ATTTCAGCGC
1401 GGAAGCGTCT GCGGAAATCG CTTCACTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTGAGCGA CAACGCCAAA
1501 CTGTTGGGCT GCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCGT
1551 TTCCGACGGC GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCCCT GCCCAAAGTT CCGCTGCGCG AGAAAATCGA TGTGGACGCA
1651 GAAGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 556; ORF 141>:

m141.pep

```

1  MSFKTDAEIA QSSTMRPIGE IAAKLGLNAD NIEPYGHYKA KINPAEAFKL
51  PQKQGRILIV TAINPTPAGE GKTTVTIGLA DALRHIGKDA VIALREPSLG
101 PVFVKGGAA GGGYAQVLPD EDINLHFTGD FHAIGAANNL LAAMLNDHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPVGVDM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRVSDADAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVNAIE SQTNNFGFAY DVELGIKDKI
451 RAIQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLEPKV PAAEKIDVDA
551 EGVINHLF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 141 shows 97.5% identity over a 558 aa overlap with a predicted ORF (ORF 141.ng) from *N. gonorrhoeae*:

m141/g141

	10	20	30	40	50	60
m141.pep	MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKL PQKQGRILIV					
g141	MSFKTDAETAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKL PQKQGRILIV					
	10	20	30	40	50	60

```

201 GCGGGGCGAA GGTAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
251 GCCATATCGG CAAAGACTCT GTGATTGCTT TCGCGGAGCC TTCTTTGGGT
301 CCGGTGTTTC GCGTGAAAGG CGGCGCGGCA GCGGCGGCT ATGCCCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTACGCCCA
401 TCGGTGCGGC AAATAATCTG CTTGCCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GCGGCGCGST
501 GGTGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGCA
551 AGCCTGTTGA CCGCGTGATG CGTCCTGACG GTTTCGATAT TACCGTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTTTG GGCAACATCC TTGTCGCTA CGCCAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTTGAAA GCGAATGSCG CGATGGCGGC ATTGCTTAAA
751 GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCCGCCTT
801 CGTACACGGC GCGCCGTTTC CCAACATCGC CCACGGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CTTGCCCGGT TTGAAACCTG ATGCGGCTGT TGTGTTGGCG ACTGTCCGCG
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAT
1051 TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTGCGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1151 TGTCGACTC CGATGCCGAG TTGCGGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251 CCGCGCGGAT TTGGCGCGCA AAGTCGTCAA TCATTGAA AGTCAAACCA
1301 ATAACCTCGG TTTCCGCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCATGTC CCAAAAAAGT GTACGGCGCG GAAGATGTTG ATTTACGCGC
1401 GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501 CTGTGGGGCT GCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCCT
1551 TTCCGAGGCG GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCGT GCCCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
1651 GAAGCGTGA TTCACGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 558; ORF 141.a>:

```

a141.pep
1 MSFKTDAEIA QSTMRPIGE IAAKLGGLND NIEPYGHYKA KINPAEAFKL
51 PQKQGRILIV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFGVKGGAA GGGYAQVLP EDINLHFTGD FHAIGAANL LAAMLNDNHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNID GMKRPVDGVM RPDGEDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WKGKGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
451 RAIQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
551 EGVIHGLF*

```

m141/a141 99.5% identity in 558 aa overlap

```

          10      20      30      40      50      60
m141.pep MSFKTDAEIAQSSTMRPIGEIAAKLGGLNDNIEPYGHYKAKINPAEAFKL PQKQGRILIV
          |||
a141      MSFKTDAEIAQSSTMRPIGEIAAKLGGLNDNIEPYGHYKAKINPAEAFKL PQKQGRILIV
          10      20      30      40      50      60

          70      80      90      100     110     120
m141.pep TAINPTPAGEGKT TVTIGLADALRHIGKDAVIALREPSLG PVFGVKGGAA GGGYAQVLP
          |||
a141      TAINPTPAGEGKT TVTIGLADALRHIGKDSVIALREPSLG PVFGVKGGAA GGGYAQVLP
          70      80      90      100     110     120

          130     140     150     160     170     180
m141.pep EDINLHFTGDFHAIGAANLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDNDRQLRNID
          |||
a141      EDINLHFTGDFHAIGAANLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDNDRQLRNID
          130     140     150     160     170     180

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 561>:

```
m142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCGTGC A3GTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCTCGTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAAAC GCTCTTCGGG CGCAGGCGTG GTCACGCCGG CAAAATGTTC
151 GGCAACATCC TGATGTTTCG CCGACAGCGT ATTGATGCAG AGGCTGCGGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
251 ACGGTGCGCG GCTCGTCGCT AACCGGCGCG ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGCTCGTAT GACATGAACG CCTGCCGCGC
351 AAGATGCCAT CGCATCACGG ATCGCAAGTT GAAAATTTTT CTGCAAAATCC
401 GCCATTTTTC CCTTTTAAAC TGTCCCGTAT ATAAGAAATG TGCACACAAG
451 GCATCCCCcC ATGTGCAGCA GTTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 562; ORF 142>:

m142.pep  
1 MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF  
51 GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLLVG NRRDRRHCHNA  
101 VTPCRVTCRD DMNACRARCH RITERSLKIF LQIRHFSPLN CPLYKNAAHK  
151 ASPHVQQF\*

Computer analysis of this amino acid sequence gave the following results:

### Homology with a predicted ORF from *N.gonorrhoeae*

ORF 142 shows 93.7% identity over a 158 aa overlap with a predicted ORF (ORF 142.ng) from *N. gonorrhoeae*:

```
m142/g142
```

	10	20	30	40	50	60
m142.pep	MRADFMFADNMPVQVRQRALYFKLSRFAAMPDVGKPLFGRRAGQGPGKMFGNILMFVRQR					
	:					
g142	MRADFMFADNMPVQVRQRAFVFYFKLSRFAAMPNMVGKPLFGRRAGQGPGKMFGNILMFVRQH					
	10	20	30	40	50	60
	70	80	90	100	110	120
m142.pep	IDAEAAVFQRDRNDSRTPVDAQHHGRRVLGNRRDRRHCAVTPCRTCVRDDMNACRARCH					
g142	IDAEAAVFQRDRNDSRTPVYAQQHGGRRVLGNRRNRHRCAVTPCRTCVRDDMNACTGTCH					
	70	80	90	100	110	120
	130	140	150	159		
m142.pep	RITERSLKI FLQIRHFSP LNCPLYKNAAHKASPHVQQFX					
g142	RITERSLKSF LQIRHFSP LNRP LYKNAAHKASPHVQQFX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 563>:

```

a142.seq
1  ATGCGTGGCG  ATTTCATGTT  TGCCGACAAT  ATGCCCGTGC  AGGTGCGCCA
51  AGCGGCCCTC  TATTTCAAGT  TGTCGGTTT  TGC CGCGATG  CCAGATGTGG
101 TAGGCAAACC  GCTCTTCGGG  CGACAGGCG  GTACGCCCGG  CAAATGTTC
151  GGCACATCC  TGA TGTCGT  CCGCGAGCG  ATTGATGAC  AGGCTGCCGT
201  TTTCCGACG  GATCGGAAT  ATTCGCGCA  TCCGGTTGAT  GCACGACATC
251  ACGGTCGGG  GCTCTCCGT  AATCCGGCA  ACCGCCGTC  TTGTAATGCC
301  GTAACGCCCT  GCCGACCGT  CTGTCGTGAT  GACATGAAC  CCTGCCGCA
351  AGGATGCCAT  CGCATCACG  AACAAGTTT  GAAAGATTTT  TC GCAATTC
401  GCCATTTTC  CCCTTAAAC  TGTCCCTAT  ATAAGAATGC  TGCACAACAG
451  GCACCCCCCA  TGTGCAGCAG  TTCTGATTCA  AAAAGCCGTC  GGTGCGACAT
501  TTCCCGCGGT  TACGGCTAT  TACGAGTTCA  ACGCATCTC  GATTTTGGCA
551  AGTTC TGCCA  ACAGGTCATT  AAGCAGCAG  ATTTTCTCG  GGCCGACGAC
601  TTCTCGATA  CGGTCGTAAC  GCTCGTCCAC  TTCTCGCGG  ATTTCTCAT
651  ACAGCTGCT  GCCCTCGGCA  GTCAGTTCA  GAAAAACAG  TCGTTGGTGC
701  TTGGAAGGTT  TCAGGCGGAC  AACCAAACCC  GCTTTTCCG  GCGGGGTGAG
751  GATACCGGTC  AGGCTGGGG  GCAAAATGCA  CGCTTGATT  GCCCAATCTT

```

405

1201 CAGGCAACCA TGTCTTGGT TGCAGGCGCA GTCTTGCTGC TGGGAGCCTT  
 1251 CTCAGTCTGT CTGATTAAAG AGATCCACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 566; ORF 143.ng>:

g143.pep  
 1 MLSEGYLGVO TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP  
 51 IVGYSDRTW KPRLGRRRLP YLLYGTIAV IVMILMPNSG SFGFGYASLA  
 101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKSAY GIOSFLANTD  
 151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL IITSFTISK  
 201 VKEYDPETYA RYHGIDVAAN QEKANWFELL KTA PKVFWTV TPVQFFCWFA  
 251 FRYMWTYSAG AIAENVWHTT DASSVGHQEA GNRYGVLA AV \*SVAAVICSF  
 301 ILAKVPNKYH KAGYFGCLAL GALGFFSIFV IYNOYALVLS YLIGIAWAG  
 351 IITYPLTIVA NALSGKHMDT YLGLFNGSVC MPQIVASLLS FVLFPMLGGH  
 401 QATMFLVAGA VLLLGAFSVC LIKEIHGGV\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 567>:

m143.seq  
 1 ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAAG  
 51 CTGCAAAATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAAAT  
 101 TGGGCTGGTT TTTCATCCTG CCGCGCTGG CGGGATGCT GGTGCAGCCG  
 151 ATTGTCGGCC ATTACTCCGA CCGCACTTGG AAGCCGCGTT TGGCGGCGCG  
 201 CCGTCTGCCG TATCTGCTTT ATGGCAGCCT GATTGCGGTT ATTGTGATGA  
 251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG  
 301 GCTTTGTCTG TCGGCGCGCT GATGATTGCG CTGTTAGACG TGTGFTCAAA  
 351 TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCAACGAGG  
 401 AGCAGAAAGG CTACGCCTAC GGGATTCAAA GTTCTTAGC AAATACGGGC  
 451 GCGGTCGTGG CGGCGATTCT GCCGTTTGTG TTTGCGTATA TCGGTTTGGC  
 501 GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTT  
 551 ATGTGGGTGC GCGCTTGCTG GTGATTACCA GCGCGTTCAC GATTTTCAAA  
 601 GTGAAGGAAT ACGATCCGGA AACCTACGCC CGTACCACG GCATCGATGT  
 651 CGCCGCGAAT CAGGAAAAG CCAACTGGAT CGAACTCTG AAAACCGCGC  
 701 CTAAGCGGTT TTGGACGGTT ACTTTGGTGC AATTCTTCTG CTGGTTCCGC  
 751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGCTCTG  
 801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT  
 851 ACGGCGTTT GCGCGCGGTG CAGTCGGTTG CGGCGGTGAT TTGTTCGTT  
 901 GTATTGCGCA AAGTGCCGAA TAAATACCAT AAGGCGGGT ATTTCGGCTG  
 951 TTTGGCTTTG GCGCGCTCG GCTTTTCTC CGTTTCTTC ATCGGCAACC  
 1001 AATACGCGCT GGTGTGTCT TATACCTAA TCGGCATCGC TTGGGCGGGC  
 1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA  
 1101 TATGGGCACT TACTTGGGCT TGTTTAACGG CTCTATCTGT ATGCCTCAAA  
 1151 TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCTATGCT GGGCGGCTTG  
 1201 CAGGCCACTA TGTCTTGGT AGGGGCGGTC GTCCTGCTGC TGGGCGCGTT  
 1251 TTCCGTGTTT CTGATTAAAG AAACACACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 568; ORF 143>:

m143.pep  
 1 MLSEGYLGVO TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP  
 51 IVGHYSRTW KPRLGRRRLP YLLYGTIAV IVMILMPNSG SFGFGYASLA  
 101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIOSFLANTG  
 151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSFTIFK  
 201 VKEYDPETYA RYHGIDVAAN QEKANWIELL KTA PKAFWTV TLVQFFCWFA  
 251 FQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AV QSVAAVICSF  
 301 VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IYNOYALVLS YTLIGIAWAG  
 351 IITYPLTIVT NALSGKHMT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL  
 401 QATMFLVGGV VLLLGAFSVF LIKEIHGGV\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m143 / g143 93.9% identity in 429 aa overlap

10 20 30 40 50 60

407

```

751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT
851 ACGGCGTTTT GCGGCGCGTG CAGTCGGTTG CCGCGGIGAT TTGTTCTGTT
901 GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGCGGGGTT ATTTCCGCTG
951 TTTGGCTTTG GCGGCGCTCG GCTTTTCTTC CGTTTCTTC ATCGGCAACC
1001 AATACGCGCT GGTGTTGTCT TAACCTTAA TCGGCATCGC TTGGGCGGGC
1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101 TATGGGCACT TACTTGGGCC TGTTAACGG CTCTATCTGT ATGCCGCAAA
1151 TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201 CAGGCCACTA TGTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
1251 TTCCGTGTTT CTGATTAAAG AAACACACGG CGGGGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 570; ORF 143.a>:

a143.pep

```

1  MLSFGLGVQ TAFTLQSSQM SRIFQTLGAD PHSLGWFFIL PPLAGMLVQP
51  IVGHYSDRTW KPRLGRRRLP YLLYGTLIAV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIOSFLANTG
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSFTIFK
201 VKEYNPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
251 FQYMWYTSAG AIAENVWHTT DASSVGYPEA GNWYGVLAAG QSVAAVICSF
301 VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
351 IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL
401 QATMFLVGGV VLLGAFSVF LIKETHGGV*

```

m143/a143 99.5% identity in 429 aa overlap

	10	20	30	40	50	60
m143.pep	MLSFGLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW					
a143	MLSFGLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW					
	10	20	30	40	50	60
m143.pep	70	80	90	100	110	120
	KPRLGRRRLPYLLYGTLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
a143	KPRLGRRRLPYLLYGTLIAVIVMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
	70	80	90	100	110	120
m143.pep	130	140	150	160	170	180
	QPFKMMVGDMVNEEQKGYAYGIGIOSFLANTGAVVAAILPFVFAFAYIGLANTAEGKVPQTVV					
a143	QPFKMMVGDMVNEEQKGYAYGIGIOSFLANTGAVVAAILPFVFAFAYIGLANTAEGKVPQTVV					
	130	140	150	160	170	180
m143.pep	190	200	210	220	230	240
	VAFYVGAALLVITSFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
a143	VAFYVGAALLVITSFTIFKVKEYNPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
	190	200	210	220	230	240
m143.pep	250	260	270	280	290	300
	TLVQFFCWFAFQYMWYTSAGAIENVWHTTDASSVGYPEAGNWNWYGVLAAGQSVAAVICSF					
a143	TLVQFFCWFAFQYMWYTSAGAIENVWHTTDASSVGYPEAGNWNWYGVLAAGQSVAAVICSF					
	250	260	270	280	290	300
m143.pep	310	320	330	340	350	360
	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
a143	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
	310	320	330	340	350	360
m143.pep	370	380	390	400	410	420
	NALSGKHMGTLYLGLFNGSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLGAFSVF					

Homology with a predicted ORF from *N. gonorrhoeae*

m144 / g144 91.3% identity in 218 aa overlap

m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFVSLADGVRENLVVSFDD
g144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTCVFVLDLGGIVQEFVSLADGVRENVPVVSFDD
m144.pep	AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
g144	AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
m144.pep	AADGRSVVLRSLATVGRRLSQRFGFG FLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
g144	AAD-----GRRLSQRFG--YFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
m144.pep	AAGRGPARGSAYSAGRTYAGRCRKRTARLNGFRPRRSIX
g144	AAGRGPARGSAYSAGRTYSGRCRKRTARLNGFRPRRSIX

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 575>:

```

a144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGCGTG
151 CGCGAAAACC TCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAGC AGATAGGCGC CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CGGTTTCGAC GGGCTGGCCG TTACCCGTTT
351 CAACGCGGTG GCGGCAGACG GCCGTTCCGT GGTGCTGCGC AGCCGCCTG.
401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTGG ATATTTCCTA
451 CCGCTTGGAC GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
501 TCGGCGACAC GGTGTTTCGAC CCGACGCTGC ACATTACTG GCGGCTGGAC
551 GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATTCC
601 GGCCGATGCC GAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
651 TATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 576; ORF 144.a>:

```

a144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFVSLADGV
51  RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLXTVGRRL SQRFGFGYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAHLL AAGRGPARGS SAYSAGRTYS
201 GRCRKRTARLN GFRRPRSI*

m144/a144 99.1% identity in 218 aa overlap

m144.pep
MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFVSLADGVRENLVVSFDD
|||||
a144
MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFVSLADGVRENLVVSFDD
|||||

m144.pep
AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV

```

Homology with a predicted ORF from *N. gonorrhoeae*

m146 / g146 90.1% identity in 212 aa overlap

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVVIDHDKVKQYGLLDFMPCLRQPPLDNFPTVRPASVEARGKYVERRRQDK					
g146	MKQIPLRLQVVIDHDKVEQYGLFDFMPCLRQPPLDNFPTVRPAFPEARAKGHVERRRQDK					
	10	20	30	40	50	60
	70	80	90	100	110	120
m146.pep	DADGFGQVRVANLRRALNVDFQNHVIACRRQRIHTLRACAVIVAKYVGVFGKSFRLDRKRLK					
	:   :					
g146	DTDSFRQVRVANLRRALNVDFQNHVIACRRQRIHALRACAVIVAEYVCVFGKSLLRDKRFLK					
	70	80	90	100	110	120
	130	140	150	160	170	180
m146.pep	LFFGNKVIMYAVCFATFRRARRVRHGNAGTVMVCQQPRHQRFARAGSGRNDKDVAFSIS					
g146	LFFGNKVIMYAVCFATFRRARRMRHGNAGTVMVCQQPRHQRFARAGSGRNDKDVAFSIS					
	130	140	150	160	170	180
	190	200	210			
m146.pep	GHIFYLYIFQPIVSQWTPSFLFADAHILPLLEX					
g146	GHIFYLYIFQPIVSQRTPYFIFADAHILPLLEX					
	190	200	210			

al46.seq

```

a146.seq
1    ATGCGCGCAA   TCCTCCTCCG   CCCGCGCCAA   GTCATCATTG   ACCACGACAA
51   AATCGAACAA   TACGGACTGT   TCGATTTCAT   GCCTTGCCTT   CGACAGCCCT
101  CTTTGATATA   TCTCCGACT   TCGCTGCCG   CGTCCGTGTA   GACGCGCAGC
151  AAGCACATCG   AAGACGAGCG   GCAGGATAAA   GATGCCGAGC   GCTTCGGGCA
201  CGCGACTCTCG   AACCTGAGCC   GCGCCCTGAA   CGTCGATTGC   CAAATACACG
251  TCATAACCTG   CCGCGCCCAA   CGCATTCACA   CCGTCCGCGC   TTGTGCGGTA
301  ATAGTTGCCG   AACACGTCG   CGTATTCCAA   AAAAGCCTCC   TGC GCGATA
351  GCGACTCAAA   CTCCTCTTG   GAAACAAGT   GATAATGTAC   GCGCTTTGCT
401  TCGCCTTCAC   CGCGCGGACG   CGTCGTGTGC   GACACGGAAA   CGCGCAACCC
451  GPTATGGTTT   GCCAACAGCC   CCGTCAACAG   CGTGGTTTTC   CCGGTGCGGG
501  AAGCGGCCGA   AATGATAAAG   ATGTTGCCTT   TTCGATAAGC   GGACATATTT
551  TTTACCTGTA   TATTTTCCAG   CGGATTGTAT   CACACAGGAC   ACCCGGTTTC
601  CTATTTGCCG   ATGCCCATAT   TTTGCCGCTA   TTGTTTGA

```

a146.pap

```

a146.pep
1  MAQILLRPRQ  VIIDHDKIEQ  YGLDFMPCL  RQPLDNFPT  VRPASVETR
51  KHIERRQDK  DADGGQRIS  NLSRLNVDF  QNHVITCRQ  RIHTLRACAV
101 IVAEHVRVFQ  KSLLRDKRLK  LFFGNKVIMY  AVCFAFTRRT  RRVRHGNAQT
151 VMVQQQPRHQ  RFARAGSGR  NDKDVAFSIS  GHIFYLYIFQ  PIVSQRTPGF
201 LFDADHILLPL  LG*

```

m146/a146 90.6% identity in 212 aa overlap

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVVIDHDKVKQYGLLDFMPCLRQPPLDNFPTVRPASVEARKGYVERRRQDK					
	:::    :::    :     :     :     :     :     :     :					
a146	MAQILLRPQVVIIDHDKIEQYGLLDFMPCLRQPPLDNFPTVRPASVETRSGKHIERRRQDK					
	10	20	30	40	50	60
	70	80	90	100	110	120

```

1101 AGGTCGCTCTG AAAGGCAGCT GGGGCGTGCA ATATTTACAA CAAAAATCCA
1151 GTGCTTTATC TGCCATATCC GAAGCGGTTA AACAAACCGAT GCTGCTTGAC
1201 AACAAAGTGC AACATTACAG CTTTTTCGGT GTAGAACAGG CAAACTGGGA
1251 CAACTTCACG CTTGAAGGAG GCGTACGCGT GGAAAAACAA AAAGCCTCCA
1301 TTCAGTACGA CAAAGCATTG ATGATCGGG AAAACTACTA CAACCACCCC
1351 CTGCCCCGACC TCGGCGCGCA CCGCCAAACC GCCCGCTCAT TCGCACTTTC
1401 GGGCAACTGG TATTTCACGC CACAACACAA ACTCAGCCTG ACCGCCTCCC
1451 ATCAGGAACG CCTGCCGTCA ACGCAAGAGC TGTACGCACA CGGCAAAACAC
1501 GTCGCCACCA ACACCTTTGA AGTCGGCAAC AAACACCTCA ACAAGAGCG
1551 TTCCAACAAT ATCGAACTCG CGCTGGGCTA CGAAGGCGAC CGCTGGCAAT
1601 ACAATCTGGC ACTCTACCGC AACCGCTTCG GTAACATCAT TTACGCCCAA
1651 ACCTTAAACG ACGGACGCGG CCCCAAATCC ATCGAAGACG ACAGCGAAAT
1701 GAAGCTCGTG CGCTACAACC AATCCGGCGC CGACTTCTAC GGCGCGGAAG
1751 GCGAAATCTA CTTCAAACCG ACACCGCGCT ACCGCATCGG CGTTTCCGGC
1801 GACTATGTAC GAGGCGGTCT GAAAAACCTG CCTTCCCTAC CCGCAGAGAG
1851 AGATGCCTAC GGCAACCGTC CTTTCATCGC ACAGGACGAC CAAATGCCCC
1901 CCCGTGTTCC GGCTGCGCGC CTCGGCTTCC ACCTGAAAGC CTCGCTGACC
1951 GACCGTATCG ATGCCAATTG GGAATCTAC CGCGTGTTCG CCCAAAAACA
2001 ACTCGCCCGC TACGAAACGC GCACGCCCGG ACACCATATG CTCACCTCG
2051 GCGCAAACTA CCGCCGCAAT ACGCGCTATG GCGAGTGGAA TTGTACGTC
2101 AAAGCCGACA ACCTGCTCAA CCAATCCGTT TACGCCACCA GCAGCTTCT
2151 CTCTGATACG CCGCAAATGG GCCGAGCTT TACCGCGGCG GTGAACGTGA
2201 AGTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>:

```

m147.pep (partial)
1 ..PHKTEQSVDL ETVSVVGKSR PRATSGLLHT STASDKIISG DTLRQKAVNL
51 GDALDGVPGI HASQYGGGAS APVIRGQTGR RIKVLNHHGE TGDMADFSFD
101 HAIMVDTALS QQVEILRGPV TLLYSSGNVA GLVDVADGKI PEKMPENGVS
151 GELGLRLSSG NLEKLTSGGI NIGLGNFVL HTEGLYRKSG DYAVPRYRNL
201 KRLPDSHADS QTGSIGLSWV GEKGFIGVAY SDRRDQYGLP AHSHEYDDCH
251 ADIIWQKSLI NKRYLQLYPH LLTEEDIDYD NPGLSCGFHD DDNAHAHHTS
301 GRPWIDLRNK RYELRAEWKQ PFPGFALRV HLNRRNDYRHD EKAGDAVENF
351 FNNQTQNARI ELRHQPIGRL KGSWGQYILQ QKSSALSALS EAVKQPMMLD
401 NKVQHYSEFFG VEQANWDNFT LEGGVRVEKQ KASIQYDKAL IDRENYNHP
451 LPDLGAHQRT ARSFALSGNW YETPQHKLSL TASHQERLPS TQELYAHGKH
501 VATNTFEVGN KHLNKRSEN IELALGYEGD RWQYNLALYR NRGNYIYAQ
551 TLNDGRGPKS IEDDSEMKLV RYNQSGADFY GAEGEIFYFKP TPRYRIGVSG
601 DYVRGRKKNL PSLPGREDAY GNRPFIAQDD QNAPRVPAAR LGFHLKASLT
651 DRIDANLDYY RVFAQNKLAR YETRTPGHHM LNLGANYRRN TRYGEWNWYV
701 KADNLLNQSV YAHSSFLSDT PQMGRSFTGG VNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

```

m147 / g147 92.3% identity in 142 aa overlap

                                10      20      30
m147.pep                      PHKTEQSVDL ETVSVVGKSR PRATSGLLHTS
                                | : | | | | | | | | | | | | | | | | | | | |
g147                          MRREAKMAQITLKPVLISILLINTPLLAQAHETEQSVGLETVSVVGKSRPRATSGLLHTS
                                10      20      30      40      50      60

                                40      50      60      70      80      90
m147.pep                      TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g147                          TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
                                70      80      90      100     110     120

                                100     110     120     130     140     150
m147.pep                      GDMADFSFDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVADGKIPEKMPENGVS
                                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g147                          GDMADFSFDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGAGQCCRWNKPPKNA

```



601 YNQSGADFYG AEGEIYFKPT PRYRIGVSGD YVRGRLKNLP SLPGREDAYG  
 651 NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY  
 701 ETRTPGHML NLGANYRRNT RYGEWNWYVK ADNLLNQSVY AHSSFLSDTP  
 751 QMGRSFTGGV NVKF\*

m147/a147 98.1% identity in 734 aa overlap

m147.pep			10	20	30
			PHKTEQSVDETQSVVVGKSRPRATSGLLHTS		
a147	MRREAKMAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETQSVVVGKSRPRATSGLLHTS				
		10	20	30	40
m147.pep		40	50	60	70
	TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRRIKVLNHHGET				
a147	TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRRIKVLNHHGET				
		70	80	90	100
m147.pep		100	110	120	130
	GDMADFSPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVADVADGKIPEKMPENGVS				
a147	GDMADFSPDHAIMVDSALSQQVEILRGPVTLTYSSGNVAGLVADVADGKIPEKMPENGVS				
		130	140	150	160
m147.pep		160	170	180	190
	ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQ				
a147	ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQ				
		190	200	210	220
m147.pep		220	230	240	250
	TGSIGLSWVGEGKGFIVAYSDDRDQYGLPAHSHEYDDCHADIWQKSLINKRYLQLYPHL				
a147	TGSIGLSWVGEGKGFIVAYSDDRDQYGLPAHSHEYDDCHADIWQKSLINKRYLQLYPHL				
		250	260	270	280
m147.pep		280	290	300	310
	LTEEDIDYDNPGLSGCFHDDDDNAHAHHSGRPWIDLRNKRYELRAEWKQFPFGFEALRVH				
a147	LTEEDIDYDNPGLSGCFHDDDDNAHAHNGKPWIDLRNKRYELRAEWKQFPFGFEALRVH				
		310	320	330	340
m147.pep		340	350	360	370
	LNRNDYRHDEKAGDAVENFFNNQTQNAIELRHQPIGRLKGSWGVQYLQKSSALSATSE				
a147	LNRNDYRHDEKAGDAVENFFNNQTQNAIELRHQPIGRLKGSWGVQYLQKSSALSATSE				
		370	380	390	400
m147.pep		400	410	420	430
	AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENNYHPL				
a147	AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENNYHPL				
		430	440	450	460
m147.pep		460	470	480	490
	PDLGAHRQTARSFALSGNWFYTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK				
a147	PDLGAHRQTARSFALSGNWFYTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK				
		490	500	510	520
m147.pep		520	530	540	550
	HLNKERSNNIELALGYEGDRWQYNLALYRNRFNGNYIYAQTLNDGRGPKSIEDDSEMKLVR				
a147	HLNKERSNNIELALGYEGDRWQYNLALYRNRFNGNYIYAQTLNDGRGPKSIEDDSEMKLVR				

101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LIATGGTMLA  
 151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

**m148 / g148** 99.0% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
g148	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m148.pep	LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA					
g148	LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m148.pep	AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGG EIVEAAAILEFTDLQGGKNI					
g148	AVEIHTDAVKPGSRVLLVDDLVATGGTMLAGLELIRKLGG EIVEAAAILEFTDLQGGKNI					
	130	140	150	160	170	180
	190	200				
m148.pep	RASGAPLFTLLQNEGCMKGX					
g148	RASGAPLFTLLQNEGCMKGX					
	190	200				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 593>:

a148.seq

1	ATGGCGTTAA	AAACATCAAA	CTTGAACAC	GCAATGCTGG	TTCATCCCGA
51	AGCTATGAGT	GTCCGTGCGC	TTGCCGACAA	AATCCGCAAA	ATCGAAAACT
101	GGCCGCAAAA	AGGCATCTTA	TTCCACGACA	TCACGCCCGT	CCTGCAAAAGC
151	GCGGAATACT	TCCGACTTTT	GGTTGATTTA	TTGGTTTACC	GCTATATGGA
201	TCAGAAAATC	GACATCGTTG	CCGGTTTGGA	CGCGCGCGGC	TTCATTATCG
251	GCGCGGCACT	CGCCTACCA	CTCAACGTCG	GTTTCGTCCC	CATCCGCAAA
301	AAAGGCAAGC	TGCCTTTTGA	AACCGTATCG	CAAAGCTACG	CGCTCGAATA
351	CGGGGAAGCT	GCGGTGGAAA	TCCACACCGA	TGCCGTCAAA	CTCGGTTGCG
401	GCGTGCTGCT	GGTCGATGAT	TTGGTTGCCA	CGGGCGGCAC	GATGCTTGCC
451	GGACTGGAGC	TGATCCGCAA	ACTCGGCGGG	GAAATTGTCG	AAGCCGCCGC
501	CATTTTGGA	TTTACCGACC	TTCAAGGCGG	CAAGAAATATC	CGTGCAAGCG
551	GCGCGCCCTT	ATTTACCCTG	CTTCAAAACG	AAGGCTGTAT	GAAGGGCTGA

This corresponds to the amino acid sequence <SEQ ID 594; ORF 148.a>:

a148.pep

1	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQS
51	AEYFRLLVDL	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LVNVGFVPIRK
101	KGKLPFETVS	QSYALEYGEA	AVEIHTDAVK	LGSRVLLVDD	LVATGGTMLA
151	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI	RASGAPLFTL	LQNEGCMKG*

**m148/a148** 99.5% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
a148	MALKTSNLEHAMLVHPEAMSVGALADKIRKIENWPQKGILFHDITPVLQSAEYFRLLVDL					
	10	20	30	40	50	60
	70	80	90	100	110	120

```

451 ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
501 CGACAGCGAA ATGAAGCTCG TGCCTACAA CCAATCCGGC GCCGACTTCT
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
601 GGCCTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
651 ACCCGGCAGA GAAGATGCCT ACGGCAACCG TCCTTTCATC GCACAGGACG
701 ACCAAAATGC CCCCCGTGTT CCGGCTGCGC GCCTCGGCTT CCACCTGAAA
751 GCCTCGCTGA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCCAAAC AAATCGCCC GCTACGAAAC GCGCAGCCCC GGACACCATA
851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
951 CAGCAGCTTT CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACGGCGC
1001 GCGTGAACGT GAAGTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 598; ORF 149>:

```

m149.pep
  1 MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY
  51 YNHPLPDLGA HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
 101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFNGY
 151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
 201 GVS GDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQNAPRV PAARLGPHLK
 251 ASLTDRIAN LDYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW
 301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 149 shows 95.9% identity over a 339 aa overlap with a predicted ORF (ORF 149.ng) from *N. gonorrhoeae*:

m149/g149

	10	20	30	40	50	60
m149.pep	MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYNHPLPDLGA					
g149	MLIDNNVRHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNQPLPDLGA					
	10	20	30	40	50	60
m149.pep	HRQTARFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
g149	HRQTARFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
	70	80	90	100	110	120
m149.pep	RSNNIELALGYEGDRWQYNLALYRNRFNGYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSG					
g149	RSNNIELALGYKGRWQYNLAAYRNRFNGYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSG					
	130	140	150	160	170	180
m149.pep	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRKLNLPGLPGREDAYGNRPFFIAQDDQNAPRV					
g149	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRKLNLPGLPGREDPYGKRPFIAQADQNAPRI					
	190	200	210	220	230	240
m149.pep	PAARLGPHLKASLTDRIANLDYRVFAQNKLARYETRTPGHMLNLGAN YRRNTRYGEW					
g149	PAARLGPHLKASLTDRIANLDYRVFAQNKLARYETRTPGHMLNLGAN YRRNTRYGEW					
	250	260	270	280	290	300
m149.pep	NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
g149	NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
	310	320	330	340		

g149-1.seq

1	ATGGCACACA	TCACACTCAA	ACCCATTGTT	TTATCAATTTC	TTTTAATCAA
51	CAACCCCTCT	CTCGCCCAAC	CGCATGAAAC	TGAGCAATCT	TGTGGCTTGG
101	AAACCGGTAG	CGTGTGCGC	AAAGAGCGTC	CGCGCGGCAG	TGTGGGGCTG
151	CTGCACACTT	CGACCGCCTC	CGACAAATCT	ATCTCCGGCG	ATACTTTTGC
201	CCAAAAAGCC	GTCAACTTGG	CGCGAGCTTT	GGACCGGGTA	CGGGGATCTG
251	ACGTTTCGCA	ATACGGCGTC	GGCGCATCCG	TCCTGGTATT	TGCGGCTCAA
301	ACGGGCGAGC	GGATTAAAGT	ATTGAACCAT	CACGGCGAAA	CGGGCGATAT
351	GGGGAGCTTT	TCTCCGATC	AGCCCAATTAT	GGTAGATACC	CGCTTGTCCG
401	AACAGGTTGA	ATCCCTGCGC	GGCGCGGTTA	CGCTTTGTGA	GACCTCGGGC
451	AATGTGGCGG	GGCTGGTCTA	TGTTTCCCGAT	GGAAAAATCC	CCGAAAAAAT
501	GCTTGAAAGC	GCGTATCTCG	CGGaagcgCG	ATTGGCTTTG	ATGACGCGGCA
551	ATTTAGAAAA	ACTGACATCC	CGACGGCGTA	ATATCCGACT	GGGCAAAAAAC
601	TTCTGCTCTC	ATACCGAAGC	CTTGTACCGC	AAATCGGGCG	ATTACGCCGT
651	ACCGCGTTAC	GCAAACTCTA	ACGCGCTGCC	CGACAGCCAT	GCCGATTTCG
701	AAACGGGCAG	CATCGGGCTG	TCTTGGGTGT	GGAAAAAGTT	GCTTTATCGC
751	CGGCATATCA	CGCGACCTCG	CGACCGCTAT	GGCCTGCTGT	CCCAACAGCCA
801	CGAATACGAT	GATTGCCACG	CGGCATATAT	TCGGCAAAAG	AGTTTGTATCA
851	ACAAACGCTA	TTTGCAGCTT	TATCCGCACT	TGTTGACCGA	AGAAGACATC
901	GATTACGACA	ATCCGGGCTT	GAGCTTCCGT	TTCCACGACG	GGCAGCGGTG
951	ACACGCACAC	ATCCACAACG	GCAAAACGGC	GATAGACGTA	CGCAACAACAC
1001	GCTACGAACT	CCGCGCCGAA	TGGAAGCAGC	CATTCCCCTG	TTTTGAAGCC
1051	CTCGCGGTAT	ATTTGACACG	CAATGACTAC	CACCACGAC	AAAAAGCAGG
1101	CGATGCAGTA	GAAAACCTTT	TCAACAAACA	AACACACACG	CGCCGTATCG
1151	AGTTGCGCCA	CCAACCCATA	GGCGTCTGTA	AAGGCACGTG	GGCGGTGCAA
1201	TATTTGGGAC	AAAAATCCAG	CGCGGTTTCC	CGCATTTCCT	AAACCGCTTCA
1251	ACACCCGATG	TGTATTGACA	ACAAATGTCC	CCATTACAGC	TTTTTTCGGT
1301	TAGAACAGGC	AAATTGGGAC	AACTTCAACG	TTGAAGGCGG	CGTACGCGTG
1351	GA AAAAACAA	AAGCCTCCAT	CCGGTACGAC	AAAGCATTTA	TGTATCGAGA
1401	AAACTACTAC	AACACGGCCC	TGCCGCGACT	CGCGCGGACG	CGCCAAACCG
1451	CCCGCTCGTT	CGCATCTGT	GGCAACTGGT	ATTTACAGCC	ACCAACA AAA
1501	CTAGCCTGCA	CCGCTTCCCA	TCAGGAAGTC	GTCGCTCAA	CGCAAGAATC
1551	GTACCGCATG	GGCAAGCAGC	TCGCCACCAA	CACCTTTGAA	GTCCGCAACA
1601	AACACTCTAA	CAAGAGCGTA	TTCCAACATA	TGCACTCGC	GCTGGGGTAC
1651	GAAAGCGGAC	GCTGGCAATA	CAATCTGGCA	CTGACCGCA	ACCGATTCCG
1701	CAACTACATT	TACGCCCAAA	CCTTAAACGA	CGGACGGCGC	CCCAATATCC
1751	TCGAAGACGA	CAGCGAATGT	AAAGCTGTGC	GTCACAACCA	ATCCGGTGCC
1801	GACTTCTACG	CGCGGAGTCA	GGAATCTTAC	TTCAAACGTA	CACCGCTGTA
1851	CCGCATCGGT	GTTTCCGGCG	ACTATGTACG	AGGCGCTGTG	AAAAACCTGC
1901	CGTCCCTTAC	CGGACGGGAA	GATCCCTACG	GCTAACCGCT	CTTCTATCGCA
1951	CACGCGGACG	AAAAAGCCGC	CGCATTTCCG	GCAGCGGCC	TGCGGTTCCA
2001	CCTGAAAACC	TCGCTAACCG	ACCGTATCGA	TGCCAAATTTG	GACTACTACC
2051	CGGTTGTTCG	CCAAACCAAC	CTTCGCGCGT	ACGAAACGCG	TACGCCCCGGA
2101	CACCTATATG	TCAACCTTCG	TGCAAACTAC	CGGCCCAATA	CGCGCTATGG
2151	CGAGTGAAT	TGGTACGTCA	AAGCGCAGAA	CCTGCTCAAC	CAATCCGTTT
2201	ACGCCCCACG	TAGCTTCTCT	TCGTGATACG	CGCAATATGG	CCGCAAGCTT
2251	ACCGGCGCGG	CAAACGTGTA	TTTTTAA		

g149-1.pap

1	MAQITLKPV	LSVILLTPRL	LAQAHETEQS	VGLETSVQVG	KSRPRATSGL
51	LHTSTASDKI	ISGDTLRQLA	VNLGDALDGO	PGIHASOYGG	GASAPVIRGG
101	TGRRIKVLNH	HGETGDMADF	SPDHAIMVDT	ALQSGOQVLR	GPVTLYSSGG
151	NVAGLVVDAD	GKIPKXPEN	GVSGEAGLRN	SSNGLEKITS	AGNIGLGNK
201	FLVHTEGVR	KSGDYAVPRY	NLNRKLPSDH	ADSQTSGL	SWVGEKGFIT
251	AYSDSDRRDY	FLGHAHSYED	DCHADIWQK	LSINKRLYLI	YPHLLTEEDI
301	DYDNPGLSCG	FHDGGAHAH	THNGKPWIDL	RNKRYELRAE	WKOPFPFGFA

551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNOSGA  
601 DFYGAEGEIIY FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPFFIA  
651 QDDQNAAPRVP AARLGPHLKA SLTDRIDANL DYRVFAQNK LARYETRTPG  
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF  
751 TGGVNVKF\*

m149-1/g149-1 96.2% identity in 758 aa overlap

	10	20	30	40	50	60
m149-1.pep	MAQTTLKPIVLSILLINTPLLAQAHE	TEQSV	DLET	VS	VVGKSR	PRATSGLLHTSTASDKI
g149-1	MAQITLKPIVLSILLINTPLLAQAHE	TEQSV	GLET	VS	VVGKSR	PRATSGLLHTSTASDKI
	10	20	30	40	50	60
	70	80	90	100	110	120
m149-1.pep	ISGDTLRQKAVNLGDALDGVPGI	HASQY	GGGAS	APVIR	GTGRR	IKVLNHHGETGDMADF
g149-1	ISGDTLRQKAVNLGDALDGVPGI	HASQY	GGGAS	APVIR	GTGRR	IKVLNHHGETGDMADF
	70	80	90	100	110	120
	130	140	150	160	170	180
m149-1.pep	SPDHAIMVDTALSQQVEILRGP	VTLLY	SSGNV	AGLV	DVADG	KIPEKMPENGVS
g149-1	SPDHAIMVDTALSQQVEILRGP	VTLLY	SSGNV	AGLV	DVADG	KIPEKMPENGVS
	130	140	150	160	170	180
	190	200	210	220	230	240
m149-1.pep	SSGNLEKLTSGGINIGLGKNF	VLHTE	GLYRK	SGDY	AVPRY	RNLKRLPDSHAD
g149-1	SSGNLEKLTSGGINIGLGKNF	VLHTE	GLYRK	SGDY	AVPRY	RNLKRLPDSHAD
	190	200	210	220	230	240
	250	260	270	280	290	300
m149-1.pep	SWVGEKGFIVGAYSDRRDQY	GLPAH	SHEY	DDCH	ADI	IWQKSLINKRYL
g149-1	SWVGEKGFIVGAYSDRRDQY	GLPAH	SHEY	DDCH	ADI	IWQKSLINKRYL
	250	260	270	280	290	300
	310	320	330	340	350	360
m149-1.pep	DYDNFGLSCGFHDDNAHAH	THSGR	PWIDL	RNKRY	ELRAE	WKQPPGFEALRV
g149-1	DYDNFGLSCGFHDDNAHAH	THSGR	PWIDL	RNKRY	ELRAE	WKQPPGFEALRV
	310	320	330	340	350	360
	370	380	390	400	410	420
m149-1.pep	RHDEKAGDAVENFFNNQTQ	NARIEL	RHQPI	GRKGS	WGVQY	LQKSSALS
g149-1	RHDEKAGDAVENFFNNKT	NARIEL	RHQPI	GRKGS	WGVQY	LQKSSALS
	370	380	390	400	410	420
	430	440	450	460	470	480
m149-1.pep	LLDNKVQHYSFFGVEQAN	WDNFT	LEGGV	RVEK	QKASI	QYDKALIDRE </td
g149-1	LIDNNVRHYSFFGVEQAN	WDNFT	LEGGV	RVEK	QKASI	RYDKALIDRE </td
	430	440	450	460	470	480
	490	500	510	520	530	540
m149-1.pep	RQTARSFALSGNWWYFT	PQHKLS	LTASH	QERLP	STQEL	YAHGKHVATNT
g149-1	RQTARSFALSGNWWYFT	PHHKLS	LTASH	QERLP	STQEL	YAHGKHVATNT
	490	500	510	520	530	540
	550	560	570	580	590	600
m149-1.pep	SNNIELALGYEGDRWQYN	LALYR	NRFG	NYI	YAQTL	NDGRGPKS
g149-1	SNNIELALGYEGDRWQYN	LALYR	NRFG	NYI	YAQTL	NDGRGPKS
	550	560	570	580	590	600
	610	620	630	640	650	660
m149-1.pep	DFYGAEGEIIYFKPTPRY	RIGV	SGDY	VRGRL	KNLPS	LPGRE
g149-1	DFYGAEGEIIYFKPTPRY	RIGV	SGDY	VRGRL	KNLPS	LPGRE
	610	620	630	640	650	660

```

351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTON ARIELRHQPI GRKGSWGVQ
401 YLGQKSSALS ATSEAVKQPM LLDNKVQHYS FFCVEQANWD NFTLEGGVRV
451 EKQKASIRYD KALIDRENY NHPPLDLGAH RQTARSFALS GNWYFTPOHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNOSGA
601 DFYGAEGEYI FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DAYGNRPLIA
651 QADQNAPRVP AARLGVHLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG
701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751 TGGVNVKF*

```

a149-1/m149-1 98.0% identity in 758 aa overlap

a149-1.pep	10	20	30	40	50	60
m149-1	10	20	30	40	50	60
a149-1.pep	70	80	90	100	110	120
m149-1	70	80	90	100	110	120
a149-1.pep	130	140	150	160	170	180
m149-1	130	140	150	160	170	180
a149-1.pep	190	200	210	220	230	240
m149-1	190	200	210	220	230	240
a149-1.pep	250	260	270	280	290	300
m149-1	250	260	270	280	290	300
a149-1.pep	310	320	330	340	350	360
m149-1	310	320	330	340	350	360
a149-1.pep	370	380	390	400	410	420
m149-1	370	380	390	400	410	420
a149-1.pep	430	440	450	460	470	480
m149-1	430	440	450	460	470	480
a149-1.pep	490	500	510	520	530	540
m149-1	490	500	510	520	530	540
a149-1.pep	550	560	570	580	590	600
m149-1	550	560	570	580	590	600
	610	620	630	640	650	660

m150.pep

1	MQNTNPPLPP	LPPEITOLLS	GLDAAQWAWL	SGYAWAKAGN	GASAGLPALQ
51	TALPAAEPFS	VTVLSASQTG	NAKSVADKAA	DSLEAAGIQV	SRAELKDYKA
101	KNIAGERRLL	LVTSTQGEGE	PPKEAVVLHK	LLNGKKAPKL	DKLQFAVLGL
151	GDSSYPNFCK	AGKDFDRRFE	ELGAKRLLER	VDADLDFTAS	ANAWTDNIAA
201	LLKEEAAKNR	ATPAPQTTTP	AGLQTAPDGR	YCKAAFPFAA	LLANQKITA
251	QSDKDVVRHIE	IDLSGSLDTH	LPGLDGLGVF	NDNDPALVREI	LDLGLIDPAT
301	EIQAGGKMMP	VARALSSHFE	LTQNTPAFVK	GYAAFAHYEE	LDKIIADNAV
351	LQDFVQNTPI	VDVLHRFPAS	LTAEQFIRLL	RPLAPRLYSI	SSAQAEVGDE
401	VHLTVGVVRF	EHGGRARTGG	ASGFILADRIE	EDGTVRVFVE	RNDGFRLPED
451	SRKQIVMIGS	GTVGAPFPRA	VQQAAEENAE	GKNWLIFGNP	HFARDFLYQT
501	EWQPFKADGF	LHRYDFAWSR	DQEEKIYVQD	KIREQAEGLN	QHWDQEGAHY
551	VCGDAAAKMAK	DVEAALLDVI	IGAGHLDEEG	AEFYLDMLRE	EKRYORDVY*

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 150 shows 91.3% identity over a 369 aa overlap with a predicted ORF (ORF 150.ng) from *N. gonorrhoeae*:

m150/g150

	210	220	230	240	250	260
m150.pep	LLKEEA	AKNRATPAPQTTPAGLQ	TAPDGRGYCKA	AFPFAALLANQKITARQSDKDV	RHIE	
g150			YCKADPFPAALLANQKITARQSDKDV	RHIE		
			10	20	30	
	270	280	290	300	310	320
m150.pep	IDLSGSDLHYLPGDALGVWFNDNPALVREILDLLGIDPATEIQAGGKMMPVARALSSHFE					
g150	IDLSGSDLHYLPGDALGVWFNDNPALVGEILDLLGINFATEIQAGGKTLPVASALLSHFE					
	40	50	60	70	80	90
	330	340	350	360	370	380
m150.pep	LTONTPAFVKGYAAAFAYEELDKIIADNAVLODFVONTPIVDVLHRFPASLTAEOFIRLL					

1701 GGATGTGATT ATCGGGGCGAG GACATTTGGA CGAAGAGGGC GCAGAAGAAT  
 1751 ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTATTGA

This corresponds to the amino acid sequence <SEQ ID 612; ORF 150.a>:

a150.pep  
 1 MQNTNPPLPP MPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ  
 51 TALPTAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA  
 101 KNIAGERRLL LVTSTQGEGE PPEEAVVLHK LLNGKKAPKL DKLQFAVLGL  
 151 GDSSYPNFCR AGKDFDKRFE ELGAKRLLER VDADLDFAAA ADGWTDNIAA  
 201 LLKEEAAKNR ATPAPQTTTP AGLQTAPDGR YCKADPFPAAL LANQKITAR  
 251 QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI DLLLGIDQAT  
 301 EIQAGGKTLP VASALLSHFE LTQNTPAFVK GYAPFADDDDE LDRIAADNAV  
 351 LQGFVQSTPI ADVLHRFPK LTAEQFAGLL RPLAPRLYSI SSSQAEVGDE  
 401 VHLTVGAVRF EHEGRARAGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED  
 451 SRKPIVMIGS GTGVAPFRAF VQQRAAENAE GKNWLFFGNP HFARDFLYQT  
 501 EWQQFAKDG LHRDYFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHY  
 551 VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY\*

m150/a150 94.8% identity in 599 aa overlap

	10	20	30	40	50	60
m150.pep	MQNTNPPLPPLPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQALPAEPPFS					
a150	MQNTNPPLPMPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQALPTAEPFS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m150.pep	VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLVTSTQGEGE					
a150	VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLVTSTQGEGE					
	70	80	90	100	110	120
	130	140	150	160	170	180
m150.pep	PPKEAVVLHKLNGKKAPKLDKLQFAVLGLGDSSYPNFCQAGKDFDRFEELGAKRLLER					
a150	PPEEAVVLHKLNGKKAPKLDKLQFAVLGLGDSSYPNFCRAGKDFDKRFEELGAKRLLER					
	130	140	150	160	170	180
	190	200	210	220	230	240
m150.pep	VDADLDFASANAWTDNIAALLKEEAAKNRATPAPQTTTPAGLQTAPDGRYCKAAPFPAA					
a150	VDADLDFAAAADGWTDNIAALLKEEAAKNRATPAPQTTTPAGLQTAPDGRYCKADPFPAAL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m150.pep	LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDPAT					
a150	LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDQAT					
	250	260	270	280	290	300
	310	320	330	340	350	360
m150.pep	EIQAGGKMMPVARALSSHFELTQNTPAFVKGYAAFAHYEELDKIADNAVLDQDFVQNTPI					
a150	EIQAGGKTLPVASALLSHFELTQNTPAFVKGYAPFADDDDELRIAADNAVLDQGFVQSTPI					
	310	320	330	340	350	360
	370	380	390	400	410	420
m150.pep	VDVLHRFPASLTAEQFIRLLRPLAPRLYISSAQAEVGDEVHLTVGVVRFEHEGRARTGG					
a150	ADVLHRFPKLTAEQFAGLLRPLAPRLYISSAQAEVGDEVHLTVGAVRFEHEGRARAGG					
	370	380	390	400	410	420
	430	440	450	460	470	480
m150.pep	ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVPFRAFVQQRAAENAE					
a150	ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVPFRAFVQQRAAENAE					
	430	440	450	460	470	480
	490	500	510	520	530	540
m150.pep	GKNWLIFGNPHF7RDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGLW					



501 NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGGK LTNIRASGTD  
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRMYL SELERRRHF  
 601 KLD\*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 615>:

m151.seq

```

1 ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA
51 AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAGAGAA
151 CGCGGCATCA CCATCCTCGC CAAAACACC GCCATCGATT ACGAAGGCTA
201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GCGCGCGAAG
251 TAGAGCGCGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
301 CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAGCCGTCCG
401 CTCGTCCGAG CTGGGTATC GACCAAACTT TCGAGCTGTT CGACAACTTG
451 GCGCGCACCG ACAGACAGTT GGATTCCCG ATTGTTTACG CTTCAGGGTT
501 GAGCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
551 CGCTGTTCGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
651 CGGCCCGCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
701 AAACCGTTGC CGTCATGAAC CACGATCAGC AAATCGCCCA AGGCCGCATC
751 AACCAGCTTT TGGGTTTCAA AGGTTTGGAA CGCGTGCCGC TTGAAGAAGC
801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATCGAAGAC ATCGGTATCG
851 GCGTAACCAT CACCGACAAA GACAATCCCA AAGGCCTACC GATGTTGAGC
901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGCT
951 GCGGGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGCGACCGCC
1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
1051 GATGCCGACG TGTTCGCGT ATCCGGGCGC GCGGAGCTGC ACCTGACCAT
1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CCAACTCGCC GTCGGCAAAAC
1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
1201 AACCTGACCG TGGATGTACC CGACGACAAAC CAAGGCGCGG TAATGGAAGA
1251 ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
1301 GACGCACCCG CCTCGAATAC CATATTCAG CGCGCGGCTT GATCGGTTTC
1351 CAAGGCGAAT TTATGACCCCT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCCGGC CGCCACAACG
1451 GCGTGTGTTG GTCCCAAGAG CAGGCGCAGG CAGTCECTTA CGCCTTGTTG
1501 AATCTGGAAG ACCGCGGCCG TATGTTGTA TCGCCCAACG ACAAAATCTA
1551 CGAAGGCATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
1651 GAAGCCGTTT GCCTGACCAC GCCAATCAAG CTGACGCTGG AAGGTGCGGT
1701 TGAGTTTATC GACGATGACG AACTCGTTGA AATCACGCCG CAATCCATCC
1751 GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCGCCG CCACTTTAAA
1801 AAGCTGGATT GA

```

This corresponds to the amino acid sequence <SEQ ID 616; ORF 151>:

m151.pep

```

1 MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
51 RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVVLLVDA
101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDITL KYTPAPSGSA
201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HDQQIAQGRI
251 NQLLGFKGLE RVPLEEAEG DIVIISGIED IGIGVTITDK DNPKGLPMLS
301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
401 NLTVDPDDN QGAVMEELGR RRGELTNMES DNGRTRLEY HIPARGLIGF
451 QGEFMTLTRG VGLMSHVFDG YAPVKPDMFG RHNGVLVSQE QGEAVAYALW
501 NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGGK LTNIRASGTD
551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRMYL SELERRRHF
601 KLD*

```

Computer analysis of this amino acid sequence gave the following results:

433

```

m151.pep      KLDX
              ||||
g151          KLDX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 617>:

```

a151.seq
1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA
51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCGCGCCA
101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAGAGAA
151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GCGGGCGAAG
251 TAGAGCGAGT TTTGGGGATG GTGGACTGCG TCGTCTTGT GGTGGACGCG
301 CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
351 TTTGGGGCTG AAACCGATTG TCGTCATCAA TAAAATCGAC AAACCGTCCG
401 CCCGTCCGAG CTGGGTCATC GACCAAACCT TCGAGCTGTT CGACAACCTG
451 GCGCGACCGC ACGAGCAGTT GGATTTCCTG ATGTTTATG CTTCCGCTCT
501 GTCCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
551 CGCTGTTCTG TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
651 CGGCCGCTCT GGTATCGGTC GTATCTT --- CGGACGTATC AAGCCCGGTC
701 AAGTTGTTGC CGTCATGAAC CACGATCAAC AAATCGCCCA AGGCCGCAATC
751 AACCAGCTTT TGGGTTTCAA AGGTTTAGAA CGCGTGCCGC TTGAAGAAGC
801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATTGAAGAC ATCGGCATCG
851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCTGCC GATGTTGAGC
901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCGCTT
951 GGCAGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGCGACCGCC
1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TCGCGTGGA AGATACGCCC
1051 GATGCCGACG TGTTCGCGCT ATCCGGGCGC GCGGAGCTGC ACCTGACCAT
1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
1151 CGCGCGTGTG GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
1201 AACCTGACCG TGGACGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
1251 ACTCGGCGCG CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGCTTC
1351 CAAGGCGAAT TTATGACCTT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCTGGC CGCCACAACG
1451 CGGTGCTGGT GTCCCAAGAG CAGGCGGAGG CAGTCGCTTA CGCCTTGTGG
1501 AATCTGGAAG ACCGCGGCCG TATGTTCTGA TCGCCCAACG ACAAATCTA
1551 CGAAGGTATG ATTTACGCA TCCACAGTCG CGACACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
1651 GAAGCCGTTT GCCTGACCAC GCCGATTAA GCTGACGCTG AAGGTGCGGT
1701 CGAGTTTATC GACGATGATG AGCTGGTAGA AATCACGCCG CAATCCATCC
1751 GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCATTTCAAA
1801 AAGCTAGATT GA

```

This corresponds to the amino acid sequence <SEQ ID 618; ORF 151.a>:

```

a151.pep
1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
51  RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVVLVDA
101 QEGPMPQTRF VTKKALALGL KPIVINKID KPSARPSWVI DQTFELFDNL
151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDITL KYTPAPSGSA
201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQVVAVMN HDQQIAQGRI
251 NQLLGFKGLE RVPLEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
351 DADVFRVSGR GELHLTILE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DNGNRTRELY HIPARGLIGF
451 QGEFMTLTRG VGLMSHVFDG YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
501 NLEDRGRMFV SPNDKIYEGM IIGIHSRND LVVNLKGGK LTNIRASGTD
551 EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRKRYL SELERRRHFK
601 KLD*

```

m151/a151 99.8% identity in 603 aa overlap

10

20

30

40

50

60

```

151 CTCGTATTCC GCCTCTGCTG GGGCATTG GGCAGcgATA CCGCCCGTTT
201 CTCcggTtTC GTCCGAGGTT GGGCAGGTAT ACGCGGCTAT CTGAAAAAcg
251 gCATTCCCGA ACAAtatcCAG CCCGACACA ACCCCTTGGG CGCACTgatg
301 gtcGTTGCGC TTTTGgcccgc cgtcTCATT CAagtcggcA CGGGGCTTTT
351 Tgccgccaat gaaaacacct tcagcaCCAa cggctacctc aaccatttgg
401 tttccgaaca tacgGGCAGC CTTATACGGA AAATCCACCT CAACTTTTTT
451 AAGCTGCTCG CCGTTTTTTC CGCAGTCCAC ATCGCCGCCG TCGCCGCATA
501 CCGCATATTC AAAAAGAAAA ACCTCGTCCG CCCGATGATA ACCGGCTTCA
551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCCGGCAA AGCCGCGCTT
601 GCCGCCGCAT TATCGGTTGC CGCGTTGCC GCAGCCGCCA TCCTGCTCCT
651 GTCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 620; ORF 152.ng>:

```

g152.pep
1  MKNKTKVWDF PTRLFWLLA ASLPFMWYSA KAGGDMLOWH TRVGLLVLFL
51  LVFRLCWGIW GSDTARFSRF VRGWAGIRGY LKNGIPEHIQ PGHNPLGALM
101 VVALLAAVSF QVGTGLFAAN ENTFTNGYL NHLVSEHTGS LIRKIHLNFF
151 KLLAVFSAVH IAAVAAYRIF KKKNLVRPMI TGFKYIEGKT SIRFAGKAAL
201 AAALSVAALA AAAILLLS*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 621>:

```

m152.seq
1  ATGAAAAACA AAACCAAAGT CTGGGACCTC CCCACCCGCC TTTTCCACTG
51  GCTGCTTGCC GCGTCCCTGC CCTTTATGTG GTATAGCGCG AAAGCCGGCG
101 GCGATATGCT GCAATGGCAC ACGCGCGTCG GGCTGTTCGT CCTTTCCTG
151 CTCGTATTTC GCCTCTGCTG GGGCATTG GGCAGCGATA CCGCCCGTTT
201 TTCCCGTTTC GTCCAAGGCT GGGCAGGCAT ACGCGGCTAT CTGAAAAACG
251 GTATTCCCGA ACACATCCAG CCCGGACACA ACCCCTTGGG CGCACTGATG
301 GTCGTTGCGC TTTTGGCGCG CGTGTCTTC CAAGTCGGCA CCGGGCTTTT
351 TGCCGCCGAT GAAAACACCT TCAGCACCAA CGGCTACCTC AACCATTTGG
401 TTTCCGAACA TACGGGCAGC CTTATGCGGA AAATCCACCT CAACTTTTTT
451 AAGCTGCTCG CCGTTTTTTC TGCAATCCAC ATCGCCGCCG TCGCCGCATA
501 CCGCGTATTC AAAAAGAAAA ACCTCATCCT CCCGATGATA ACCGGCTTCA
551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCAGGCAA AGCCGCGCTT
601 GCCGCCGCAT TATCGGTTGC CTCGCTTGCC GCAGCCGCCA TCCTGCTCCT
651 GTCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 622; ORF 152>:

```

m152.pep
1  MKNKTKVWDL PTRLFWLLA ASLPFMWYSA KAGGDMLOWH TRVGLFVLEL
51  LVFRLCWGIW GSDTARFSRF VQGWAGIRGY LKNGIPEHIQ PGHNPLGALM
101 VVALLAAVSF QVGTGLFAAD ENTFTNGYL NHLVSEHTGS LMRKIHLNFF
151 KLLAVFSAIH IAAVAAYRVF KKKNLILPMI TGFKYIEGKT SIRFAGKAAL
201 AAALSVASLA AAAILLLS*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 152 shows 95.4% identity over a 218 aa overlap with a predicted ORF (ORF 152.ng)

from *N. gonorrhoeae*:

m152/g152

```

          10      20      30      40      50      60
m152.pep  MKNKTKVWDL PTRLFWLLAASLPFMWYSAKAGGDMLOWHTRVGLFVLEL LVFRLCWGIW
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g152      MKNKTKVWDFPTRLFWLLAASLPFMWYSAKAGGDMLOWHTRVGLLVFLFLVLFRLCWGIW
          10      20      30      40      50      60

          70      80      90     100     110     120
m152.pep  GSDTARFSRFVQGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAD
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g152      GSDTARFSRFVRGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAN
          70      80      90     100     110     120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 625>:

```
g153.seq
1  atgggggtttg cttacAgat gacgtatc gaggtCGGGA taccggaggc
51  ggcacccgtc ctttCgtGC CCGAGATgat gcgcctgatG GTGTTcCagg
101 attATGGTTT TttggcCGAA GTGATGTTG TGctgaCTTt cGGCGgcCG
151 GTTCGTGTTc TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA
201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA
251 GGCAGGCGAT GATGGTGGAT GTGTTTTTGT TTCCACTCTT GGTGGCGTAT
301 ATCAAGCTCT CGTCTGTGGC AAAGGTTTCG TTCGGGCGCG CGTTTTATCT
351 GATGTTCCGC CTGTCGGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC
401 AGCATTGGGT GTATTTCCAA ATCGGGCGGC TGACGGGGAA TAATGCGGTT
451 CAGACGGCAT CGGAAGGCAA AACCTGTTGC AGCGCTGCC TGTATTTccg
501 cgacAGTgcc gaatcccccT GCGGGGTGTg cgCGcggaA CTgtacggcg
551 gacggccgaa aagtCTGAGt atttCgtCGG CGTTTCTgac ggcggcgTT
601 GTTTTGATT TCCctgCcaa TATCctgccg attaTGAttt cgtccAATCc
651 tgccggccag GAGGcCAACA CCATCTTAG CGGCATCGCT TATATGTGGG
701 ACGaggcgA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG
751 GTGCCGTGC TGAAGATTGC GGCAATGTCG GTTTTGATTG CCGCGGCACG
801 GTTCGCTTTG CCGCGGGCG CAAAGAAAT GTCGCACCTC tacCGCATCA
851 CCGAAGCGGT CGGCCGCTG TCGATGATTG ATATTTTGT GATTATTATT
901 TTGATGTGTT CGTTCcaca TTATGCCCGC CGCGTCATTc CGGGCAGTGC
951 GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001 ATTTGACCC GCGCTGCTT TGGGACAAAC GCGCTTCAGA CCGCATTGCT
1051 TTCAACGAAA CGGAAAAATA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 626; ORF 153.ng>:

```
g153.pep
1  MGFAYSMTYI EVGIPEAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
51  VLFLLLCLYV YAALIRKQAY PALRLATRVV VRLRQAMMVD VFFVSTLVAY
101 IKLSSVAKVR FGPAPYLMFA LSVMLIRTSV SVPQHWVYFQ IGRLTGNNAV
151 QTASBGKTCC SRCLYFRDSA ESPCGVCGAE LYGGRPKSLs ISSAFLTAaV
201 VLYFPANILP IMISSNPAT EANTIFSGIA YMWDEGDRLI AAVIFSASIL
251 VPVLKIAAMS VLIAAARFAL PAGAKKLSHL YRITEAVGRW SMIDIFVILI
301 LMCSEHTYAA RVIPGSAAYV FCLVVILTML SAYYFDRLL WDKRASDGLA
351 FNETEKYD*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 627>:

```
m153.seq
1  ATGGCGTTTG CTTACGGTAT GACGTATATC GAGGTCGGGA TACCGGGTGC
51  GGCATCCGTC CTTTCGCTGC CCGAGATGAT GCGCCTGATG GTGTTTCAGG
101 ATTATGGTTT TTGGCCGAA GTGATGTTG TGCTGACTTT CCGCGCGCCG
151 GTTCGTGTTT TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA
201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA
251 GACAGGCGAT GATGGTGGAT GTGTTTTTGT TTCCACTTT GGTGGCGTAT
301 ATCAAGCTCT CGTCTGTGGC AGAGCTTCGC TTCGGGCGCG CGTTTTATCT
351 GATGTTCCGC CTGTCACTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC
401 AGCATTGGGT GTATTTTCAA ATCGGGCGGC TGACGGGGAA TAATGCGGTT
451 CAGACGGCAT CGGAAGGTAA AACCTGTTGC AGCGCTGCC TGTATTTCCG
501 CGACAGTGCC GAATCCCCCT GCGGCGTGTG CGGTGCGGAA CTGTACCGCC
551 GACGCGCGAA AAGTCTGAGT ATTTCTGTCG CGTTTCTGAC GCGCGCGGTT
601 ATTTTGATT TCCCTGCCAA TATCCTGCCG ATTATGATT CTGCCAATCC
651 TGCCGCCACG GAGGTCAATA CCATCCTTAA CGGCATCGCT TATATGTGGG
701 ACGAGGCGCA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG
751 GTGCCGGTAC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGTCCGCCCG
801 CTTGCTTTG CCAACGGGTG CAAAGAAAT GTCGCACCTC TACCGCATCA
851 CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTGT GATTATTATT
901 TTGATGTGTT CGTTCACAC TTATGCCCGC CGCGTCATTc CGGGCAGTGC
951 GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001 ATTTGACCC GCGCTGCTT TGGGACAAAC GCGCTTCAGA CCGCATTGCT
1051 TTCAATGAAA CGGAAAAACA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 628; ORF 153>:

```
m153.pep
1  MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
51  VLFLLLCLYV YAALIRKQAY PALRLATRVV VRLRQAMMVD VFFVSTLVAY
```

439

```

851 CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTGT GATTATTATT
901 TTGATGTGTT CGTCCACAC TTATGCCGCG CGCGTCATTC CGGGCAGTGC
951 GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001 ATTCGACCC GCGCCTGCTT TGGACAAAC GCGCTTCAGA CGGCATTGCT
1051 TTCAATGAAA CGGAAAAACA TGACTGA

```

This corresponds to the amino acid sequence <SEQ ID 630; ORF 153.a>:

```

a153.pep
1  MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
51  VLFLLCLYV YAALIRKQAY PALRLATRV VRLRQAMMVD VEFVSTLVAY
101 IKLSSVAEVR FGSIFYLMFA LSVMLIRTSV SVPOHWVYFQ IGRLTGDNV
151 QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYRRRPKSL SISAFLTA
201 ILYFPANILP IMISSNPAAT EVNTILNGIA YMWDEGDR LI AAVIFSASIL
251 VPVLKIAAMS VLIASARFAL PTGAKKLSHL YRITEAVGRW SMIDIFV
301 LMCSFHTYAA RVIPGSAAVY FCLVVILTML SAYYFDPRL LWDKRASDGI
351 FNETEKHD*

```

m153/a153 99.7% identity in 358 aa overlap

	10	20	30	40	50	60
m153.pep	MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFDYGFLEAVMFVLTFGAPVLFLLCLYV					
a153	MAFAYGMTYIEVGIPGAASVLSLPEMMRLMVFDYGFLEAVMFVLTFGAPVLFLLCLYV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m153.pep	YAALIRKQAYPALRLATRVVRLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGPAFYLMFA					
a153	YAALIRKQAYPALRLATRVVRLRQAMMVDVFFVSTLVAYIKLSSVAEVRFGSAFYLMFA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m153.pep	LSVMLIRTSVSVPOHWVYFQIGRLTGDNVQTASEGKTCCSRCLYFRDSAESPCGVCGAE					
a153	LSVMLIRTSVSVPOHWVYFQIGRLTGDNVQTASEGKTCCSRCLYFRDSAESPCGVCGAE					
	130	140	150	160	170	180
	190	200	210	220	230	240
m153.pep	LYRRRPKSLSSISAFLTAAVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDR LI					
a153	LYRRRPKSLSSISAFLTAAVILYFPANILPIMISSNPAATEVNTILNGIAYMWDEGDR LI					
	190	200	210	220	230	240
	250	260	270	280	290	300
m153.pep	AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHL YRITEAVGRWSMIDIFV					
a153	AAVIFSASILVPVLKIAAMSVLIASARFALPTGAKKLSHL YRITEAVGRWSMIDIFV					
	250	260	270	280	290	300
	310	320	330	340	350	359
m153.pep	LMCSFHTYAA RVIPGSAAVYFCLVVILTMLSAYYFDPRL LWDKRASDGI AFNETEKHD					
a153	LMCSFHTYAA RVIPGSAAVYFCLVVILTMLSAYYFDPRL LWDKRASDGI AFNETEKHD					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 631>:

```

g154.seq
1  ATGACTGACA ACAGCCCTCC TCCAAACGGA CACGCTCAAG CACGCGTCCG
51  CAAAAACAAC accttctctCT CGGCCGTCTG GCTGGTCCCG CTGATCGCGC
101 TGATTGCCGG CGGCTGGCTT TGGGTAAAGG AAATCCGCAA CAGGGGCGCT
151 GTGGTTACGC TCTTGATGGA CAGCGCGGAA GGCATCGAAG TCAACAATAC
201 GGTCAATTAAG GTATTGAGCA TCGATGTCGG ACGCGTTACC CGAATCAAAC
251 TGC GCGACGA CCAAAAAGGC GTGGAAGTTA CTGCCCAACT CAATCGGCAC
301 GTATCCGGCC TCATCCGCAG CGATACCCAG TTTTGGGTGG TCAAGCCGCG
351 TATCGACCAA AGCGGcgtaa CCGGTTTGGG TACGCTGCTT TCGGGTTCGT

```

```

1201 TCACCTAAGC TGCACCGCA TACCGTTTAT GCAGGCGATA CCGTTATCGC
1251 GACCCAGGGC GCGGTTTGG ACGATTGCA GGTCAAATTG CCGGATTGCG
1301 TGGACAAGTT CGACAACTG CCTTTAGATA AGACGGTTGC CGAATTGAAC
1351 GGTTCGCTTG CCGAGCTCAA ATCCACACTC AAATCTGCCA ATGCCGCCCT
1401 AAGCTCCATC GACAACTGG TCGGCAAAAC GCAGACACAA AACATTCCGA
1451 ACGAACTGAA CCAAAACCTG AAAGAGTTGC GCACAACCTT GCAAGGCGTA
1501 TCGCCGCAAT CGCCTATCTA CGGCGACGTA CAAAATACGC TGCAAAGTTT
1551 GGACAAACTT TAAAAGACG TTCAACCCGT GATTAATACT TTGAAAGAAA
1601 AACCAACGCG GCTGATTTTC AACAGCAGCA GCAAAGACCC TATCCCGAAA
1651 GGAAGCCGAT AA

```

This corresponds to the amino acid sequence <SEQ ID 634 ORF 154.a>:

```

m154 .pep
  1  MTDNSPPPNG HAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP
 51  VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
101  VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ
151  VQDIPPVTAI GQSGRLRLNI GKNDRIILNVN SPVLYENFMV GQVESAHFDP
201  SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI KLNSAPLPAL
251  LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ
301  SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFELGW IPVRIRIEPS
351  RLEINAEQS KEHWKQFQT ALNKGLTATI SSNNLTGSK MIELNDQPSA
401  SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLDDKFDKL PLDKTVAELN
451  GSLAELKSTL KSANAALSSI DKLVGKPGTQ NIPNELNQTQ KELRTTLQGV
501  SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK
551  GSR*

```

m154 / g154 97.8% identity in 553 aa overlap

	10	20	30	40	50	60
m154 .pep	MTDNSPPPNGHAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP VVTLLMDSAE					
g154	MTDNSPPPNGHAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP VVTLLMDSAE					
	10	20	30	40	50	60
m154 .pep	70	80	90	100	110	120
	GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD VSGLIRSDTQ FWVVKPRIDQ					
g154	GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD VSGLIRSDTQ FWVVKPRIDQ					
	70	80	90	100	110	120
m154 .pep	130	140	150	160	170	180
	SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGRLRLNI GKNDRIILNVN					
g154	SGVTGLGTLT SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGRLRLNI GKNDRIILNVN					
	130	140	150	160	170	180
m154 .pep	190	200	210	220	230	240
	SPVLYENFMV GQVESAHFDP SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI					
g154	SPVLYENFMV GQIESAHFDP SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI					
	190	200	210	220	230	240
m154 .pep	250	260	270	280	290	300
	KLNSAPLPALL SG AISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ					
g154	KLNSAPLPALL SG AISFDSP KTKNSKNVKS EDSFTLYDSR SEIANLPDDR SLYYTAFFKQ					
	250	260	270	280	290	300
m154 .pep	310	320	330	340	350	360
	SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGWIP VRIRIEPS RLEINAEQS					
g154	SVRGLTVGSP VEYKGLNVGMVSDVPYFDRN DSLHLFENGWIP VRIRIEPS RLEINAEQS					
	310	320	330	340	350	360
m154 .pep	370	380	390	400	410	420
	KEHWKQFQTALNKGLTATISSNNLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG					
g154	KEHWKQFQTALNKGLTATISSNNLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATRG					

```

151 VQDIPPVTAI GQSGRLRLNLI GKNDRIILNVN SPVLYENFMV GQVESAHFDP
201 SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTSGSI KLNSAPLPAL
251 LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKQ
301 SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351 RLEINADEQS KEHWKQFQT ALNKGLTATI SSNNLLTGSK MIELNDQPSA
401 SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLLDKFDKL PLDKTVAELN
451 GSLAEKSTL KSANAALSSI DKLVGKPTQ NIPNELNQT LKELRTTLQGV
501 SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK
551 GSR*

```

## m154/a154 100.0% identity in 553 aa overlap

	10	20	30	40	50	60
m154.pep	MTDNSPPPNHGAQARVRKNNITLSAVWLVLIALIAGGWLWVKEIRNRGPVVTLLMDSAE					
a154	MTDNSPPPNHGAQARVRKNNITLSAVWLVLIALIAGGWLWVKEIRNRGPVVTLLMDSAE					
	10	20	30	40	50	60
m154.pep	70	80	90	100	110	120
	GIEVNNTVIKVLSDVGRVTRIKLRDDQKGEVTAQLNADVSGLIRSDTQFWVVKPRIDQ					
a154	GIEVNNTVIKVLSDVGRVTRIKLRDDQKGEVTAQLNADVSGLIRSDTQFWVVKPRIDQ					
	70	80	90	100	110	120
m154.pep	130	140	150	160	170	180
	SGVTGLGTLTLLSGSYIAFTPGKSDEAKDVFQVQDIPPVTAIGQSGRLRLNLIGKNDRIILNVN					
a154	SGVTGLGTLTLLSGSYIAFTPGKSDEAKDVFQVQDIPPVTAIGQSGRLRLNLIGKNDRIILNVN					
	130	140	150	160	170	180
m154.pep	190	200	210	220	230	240
	SPVLYENFMVGQVESAHFDPDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTSGSI					
a154	SPVLYENFMVGQVESAHFDPDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTSGSI					
	190	200	210	220	230	240
m154.pep	250	260	270	280	290	300
	KLNSAPLPALLSGAISFDSPKTKNSKNVKS EDSFTLYDSRSEVANLPDDRSLYYTAFFKQ					
a154	KLNSAPLPALLSGAISFDSPKTKNSKNVKS EDSFTLYDSRSEVANLPDDRSLYYTAFFKQ					
	250	260	270	280	290	300
m154.pep	310	320	330	340	350	360
	SVRGLTVGSPVEYKGLNVGVVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEQS					
a154	SVRGLTVGSPVEYKGLNVGVVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEQS					
	310	320	330	340	350	360
m154.pep	370	380	390	400	410	420
	KEHWKQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG					
a154	KEHWKQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG					
	370	380	390	400	410	420
m154.pep	430	440	450	460	470	480
	GGLDDLQVKLADLLDKFDKLPDKTVAELNGSLAEKSTLKSANAALSSIDKLVGKPTQ					
a154	GGLDDLQVKLADLLDKFDKLPDKTVAELNGSLAEKSTLKSANAALSSIDKLVGKPTQ					
	430	440	450	460	470	480
m154.pep	490	500	510	520	530	540
	NIPNELNQT LKELRTTLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF					
a154	NIPNELNQT LKELRTTLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF					
	490	500	510	520	530	540

```

301 TTGGTCAAG CCTTGCGCGC CAAGAAAGTG AACGCGCTGG CGATGGATAT
351 GGTGCCCCGC ATTTGCGCGC CGCAGGCTTT GGACGCTTTG TCTTCGATGG
401 CAAACATCAG CGGCTACCGC GCCGTAATTG AAGCCGCCAA CGCCTTCGGC
451 CGTTTCTTCA CCGGTCAAAT TACCGCGGCC GGCAAAGTGC CGCCCGCGCA
501 GGTTTTGGTG ATTTGGTGCAG GTGTGGCAGG TTTGGCGGCG ATCGGTACGG
551 CAAACTCGCT CGGCGCAGTG GTACGCGCGT TCGATACCCG CTGGGAAGTG
601 GCGGAACAAA TCGAATCGAT GGGCGGCAAG TTCCTGAAAC TCGACTTCCC
651 ACAAGAATCG GGGCGCAGCG GAGACGGCTA CGCCAAAGTG ATGAGCGACG
701 AATTATATCG AGCCGAGATG AAGCTCTTTG CCGAGCAGGC GAAAGAAGTG
751 GACATCATCA TCACCACCGC CGCCATTCCG GGCAAACCCG CGCCCAAGCT
801 GATTACCAA GAAATGGTGG AAAGCATGAA ATCCGGCTCC GTCATCGTCG
851 ATTTGGCGGC GGGCAGCGGC GGCAACTGCG AACTCACCCG CCGGGCGGAA
901 TTGTCGTAA CCGGCAACGG CGTGAAATC ATCGGCTACA CGACATGGC
951 AAACCGCCTT GCCGACAGT CTCCAGCT TTACGCCACC AACTTGGTCA
1001 ACCTGACCAA GCTGTTAAGC CCGAACAAAG ACGGCGAAAT CACGTGGAC
1051 TTCGAAGACG TGAATATCCG CAACATGACC GTTACCCACG ACGGCGAAAT
1101 CACCTTCCCG CCTCCGCCGA TTCAAGTTTC CGCCAGCCG CAGCAACCGC
1151 CGTCTGAAAA AGCCGTGCCT GCCGCCAAGC CCGAGCCAAA ACCCGTCCC
1201 CTGTGAAAAA AACTCGCGCC CGCCGTCATC GCCCGCTCT TGGTACTGTG
1251 GGTGCGCGCG GTGCGACCCG CAGCATTCTT GAACCACTTT ATCGTGTTCG
1301 TTCTCGCCTG CGTCATCGGC TACTACGTCG TCTGGAACGT CAGCCACTCG
1351 CTGCACACAC CGCTGATGTC GGTAAACCAAC GCCATCTCCG GCATCATCGT
1401 CGTCCGCGCG CTGCTGCAAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC
1451 TGTGCTTTGT TGCCATCCTG ATTGCCGGCA TCAACATCTT CGGCGGCTTT
1501 GCGGTAACAC GCGGTATGCT GAATATGTTT AAGAAAGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 640; ORF 155>:

```

m155.pep
1  MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY
51  QTAGATVADK AAVWVCP- LIY
 KVNAPSEQEL PLLNEGQTIV SFLWRQNEA
101 LVEALRAKKV NALAMDMPVR ISRAQALDAL SSMANISGYR AVIEAANAFG
151 RFFTGQITAA GKVPVPAQVLV IGAGVAGLAA IGTANSLGAV VRAFDTRLEV
201 AEQIESMGGK FLKLDFFQES GSGDGYAKV MSDEFIAEM KLFQAEQKEV
251 DIIITTAAP GKPAKPLITK EMVESMKSGS VIVDLAAATG GNCELTRPGE
301 LSVTGNGVKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKGEITLD
351 FEDVIIRNMT VTHDGEITFP PPPIQVSAQP QQTPEKAVP AAKPEPKVVP
401 LWKKLAPAVI AAVLVLVVGA VAPAAFLNHF IVFVLACVIG YVVVWNVSHS
451 LHTPLMSVTN AISGIIVVGA LLQIGQGNF VSLLSFVAIL IAGINIFGGF
501 AVTRRMLNMF KKG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 155 shows 97.9% identity over a 513 aa overlap with a predicted ORF (ORF 155.ng) from *N. gonorrhoeae*:

```

m155 / g155 97.9% identity in 513 aa overlap
              10      20      30      40      50      60
m155.pep      MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAYQTAGATVADK
              |||
g155           MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAYQTAGATVADK
              10      20      30      40      50      60

              70      80      90      100     110     120
m155.pep      AAVWVCP- LIY
KVNAPSEQELPLLNEGQTIVSFLWRQNEALVEALRAKKVNALAMDMPVR
              |||
g155           AAVWACPLIYKVNAPSEGELPLLKEGQTIVSFLWRQNEALVEALRAKKVNALAMDMPVR
              70      80      90      100     110     120

              130     140     150     160     170     180
m155.pep      ISRAQALDALSSMANISGYRAVIEAANAFGRFFTGQITAAGKVPAQVLVIGAGVAGLAA
              |||
g155           ISRAQALDALSSMANISGYRAVIEAANAFGRFFTGQITAAGKVPAQVLVIGAGVAGLAA

```



```

1201 CTGTGGAAAA AACTCGCGCC CGCCNTNATC GCCGCCGTGT TGGTACTGTG
1251 GGTCGGCGCG GTGCGACCCG CAGCATTCTT GAACCACTTT ATCGTCTTCG
1301 TCCTCGCCTG CGTCATCGGC TACTATGTCT TTTGGAACCT CAGCCACTCG
1351 CTGCACACAC CGCTGATGTC GGTGACCAAC GCCATTTCCT GCATCATCGT
1401 CGTCGGCGCG CTGCTGCAAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC
1451 TGTCGTTTGT TGCCATCCTG ATTGCCAGCA TCAACATCTT CGGCGGCTTC
1501 TTTGTAACGC GGCGGATGCT GAATATGTTT AGGAAAGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 642; ORF 155.a>:

```

a155.pep
  1 MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY
  51 QAAGATVADK AAVWAYPLIY KVNAPSEDEL PLLKEGQTIV SFLWPRQNEA
 101 LVEALRAKKV NALAMDMPVR ISRAQALDXL SXMANISGYR AVIEAANAFG
 151 RXFTGQITAA GKVPAPQVLV IGAGVAGLAA IGTANSLGAV VRVFDTRLXV
 201 AEQLESMGGK FLKLDFFQES GSGDGYAKV MSDEFIAEM KLFAEQAKEV
 251 DIIITTAAP GKPAKXXXX EMVESMKPGS VIVDLAAATG GNCELTKQGE
 301 LFTGNGVKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKDGEITLD
 351 FEDVIIRNMT VTRDGEITFP PPPIQVSAQP QQTPEKAAP AAKPEPKPVP
 401 LWKKLAPAXI AAVLVLVVGA VAPAAFLNHF IVFVLACVIG YVYVWNVSHS
 451 LHPTLMSTVN AISGIIVVGA LLQIGQGNF VSLLSFVAIL IASINIFGGF
 501 FVTRRLNMF RKG*

```

m155/a155 95.3% identity in 513 aa overlap

	10	20	30	40	50	60
m155.pep	MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAYQTAGATVADK					
a155	MKIGIPRESLSGETRVACTPATVALLGKLG FETVVESGAGLAASLDDAAYQAAGATVADK					
	10	20	30	40	50	60
	70	80	90	100	110	120
m155.pep	AAVWVCPLIYKVNAPSEQLPLLN EGQTIVSFLWPRQNEALVEALRAKKVNALAMDMPVR					
a155	AAVWAYPLIYKVNAPSEDELPLLN EGQTIVSFLWPRQNEALVEALRAKKVNALAMDMPVR					
	70	80	90	100	110	120
	130	140	150	160	170	180
m155.pep	ISRAQALDALSSMANISGYRAVIEAANAFGRFFTGQITAAGKVPPAQVLVIGAGVAGLAA					
a155	ISRAQALDXLSXMANISGYRAVIEAANAFGRXFTGQITAAGKVPPAQVLVIGAGVAGLAA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m155.pep	IGTANSLGAVVRAFDTRLXVAEQIESMGGKFLKLDFFQESGSGDGYAKVMSDEFIAAEM					
a155	IGTANSLGAVRVFDTRLXVAEQLESMGGKFLKLDFFQESGSGDGYAKVMSDEFIAAEM					
	190	200	210	220	230	240
	250	260	270	280	290	300
m155.pep	KLFAEQAKEVDIIITTAAPGKPAKKLITKEMVESMKSGSVIVDLAAATGGNCELTRPGE					
a155	KLFAEQAKEVDIIITTAAPGKPAKXXXXKEMVESMKPGSVIVDLAAATGGNCELTKQGE					
	250	260	270	280	290	300
	310	320	330	340	350	360
m155.pep	LSVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT					
a155	LFTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNMT					
	310	320	330	340	350	360
	370	380	390	400	410	420
m155.pep	VTHDGEITFPPIQVSAQPQQTPEKAAPAAKPEPKVPPLWKKLAPAVIAAVLVLVVGA					
a155	VTRDGEITFPPIQVSAQPQQTPEKAAPAAKPEPKVPPLWKKLAPAXIAAVLVLVVGA					
	370	380	390	400	410	420

449

```

g156      FAPFAAAVLTAHATGNAGQATVNTLAGLFILFRLAFIWCYIADKAALRSLMWAGGFACTV
              70      80      90      100     110     120

m156.pep  GLFVAAAX
           |||||
g156      GLFVAAAX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 647>:

```

a156.seq
1  ATGACTTTCG CCTATTGGTG TATTCTGATT GCCTACCTAT TGCCGCTTTT
51 TTGTGCGGCG TATGCCAAAA AAGCGGCGCG ATTCCGGTTT AAAGACAACC
101 ACAATCCGCG CGATTTTCTG GCGCGCACGC AAGGCACAGC CGCCCGTGCC
151 CACGCGCGCG AGCAAAACGG TTTGAAGCC TTGCACCGT TTGCAGCCGC
201 CGTTTTGACG GCACACGCAA CCGGCAATGC CGGACAAGCA ACCGTCAACA
251 CGCTTGCCCG CCTGTTTCATC CTGTTCCGCC TCGCCTTTAT CTGGTGCTAC
301 ATCGCAGACA AAGCAGCATT ACGCTCGCTG ATGTGGGTGG GCGGATTTGT
351 CTGCACCGTC GGGCTGTTTG TCGTGGCTGC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 648; ORF 156.a>:

```

a156.pep
1  MTFAYWCILI AYLLPLFCAA YAKKAGGERF KDNHNPRDFL ARTQGTAAARA
51  HAAQONGFEA FAPFAAAVLTAHATGNAGQA TVNTLAGLFI LFLAFIWCY
101 IADKAALRSL MWVGGFVCTV GLFVVAAX*

m156/a156  90.6% identity in 127 aa overlap

              10      20      30      40      50      60
m156.pep    MTFAYWCILIACLLPLFCAAYAKKAGGFRFKDNHNPRGFLAHTQGAARAHAHQNGFEA
              |||||
a156        MTFAYWCILIAYLLPLFCAAYAKKAGGFRFKDNHNPRDFLARTQGTAAARAHAHQNGFEA
              10      20      30      40      50      60

              70      80      90      100     110     120
m156.pep    FAPFAAAVLTAHATGNAAQSTINTLACLFILFRLAFIWCYIADKAAMRSLMWAGGFACTV
              |||||
a156        FAPFAAAVLTAHATGNAGQATVNTLAGLFILFRLAFIWCYIADKAALRSLMWVGGFVCTV
              70      80      90      100     110     120

m156.pep    GLFVAAAX
              |||||
a156        GLFVVAAX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 649>:

```

g157.seq
1  atgaggaacg aggAAAAACg cgcctgqgc cgcgaattgC gCgGgcgqgc
51  ttcgcAAATg GGgcgagacg tGCGggCGGC GCGGgCgatA Aaaatcaacc
101 gcctgctcaa aCGTtatatc AAGCGCggtc gGaAaatcgG CGTGtATTgg
151 cCGATGGGCA AGGAATTGcg TTTGGGCGgc tTgtcCGCG CGGCGCAAAA
201 ACGCgCGGCA AAactctatc tgccttATAT CGAACCgCAC ACGCGGCGGA
251 TGTGgTTTAC GCCGTATCCT GAACGCGGAA TGGAACGGGA ACGCAAGCGC
301 GGTAGGGCGA AGCTGCATGT CCCTCAGTTT GCAGGGCGCA AAATCCGCGT
351 GCACGGTTTG TCGGTATTGC TCGTCCCGCT TGTCGGCATA GACCGCGAAG
401 GCTACCGTTT GGGGCAGGCA GCGCGCTATT ACGATGCGAC GCTTTCGGCG
451 ATGAAATACC GTTTGCAGGC GAAAACCGTG GCGGTGGGCT TTGCCTGCCA
501 GTTGTGGGAC AGGCTCCAC GCGAGGCGCA CGACCTGCCG CTGGACGGTT
551 TTGTATCGGA AGCGGGGATA TTGTGTTTTT AG

```

451

```

101 GCCTGCTCAA ACGTTATATC AAGCGTGGTC GGAAATCGG CGTGATTGG
151 CCGATGGGCA AGGAATTGCG TTTGGACGCG TTTGTCCGCG CGGCGCAAAA
201 ACGCGGTGCA AAACCTTATC TGCCTTATAT CGAACCGCGT TCGCGGCGGA
251 TGTGGTTTAC GCCGTATCCT GAAAGCGGAA TGGAAACGGGA GCGCATACGG
301 GGCAGGGCGA AGTTGAACGT GCCGCAGTTT GCAGGGCGCA AAATCCGCGT
351 GCACGCTTTC TCGGTATTGC TCGTCCCGCT TGTGGGCATA GACCGCGAGG
401 GCTACCGCTT AGGACAGGCA GGC GGCTATT ACATGCGAC GCTTGGCGCG
451 ATGAAATACC GTTTGCAGGC AAAAACCGTG GCGTGGGCT TTGCCTGCCA
501 GTTTGTGGAC AGGCTGCCGC GCGAACCGCA CGATCTGCTG CTGGACGGTT
551 TTGTGTCGGA GCGGGGATA TTGTGCTTTT AG

```

This corresponds to the amino acid sequence <SEQ ID 654; ORF 157.a>:

**a157.pep**

```

1 MRNEEKHALR RELRRARAQM GHQGRLAAGQ TINRLLKRYI KRGRKIGVYW
51 PMGKELRLDG FVRAAQKRGK KLYLPYIEPR SRRMWFTTYP ESGMERERIR
101 GRAKLNVPQF AGRKIRVHGL SVLLVPLVGI DREGYRLGQA GGYDTLAA
151 MKYRLQAKTV GVGFAQQFVD RLPREPHDLL LDGFVSEAGI LCF*

```

**m157/a157** 82.4% identity in 193 aa overlap

	10	20	30	40	50	60
m157.pep	MRNEEKRALRRELRRRSQMGRDVR	AAATVKINHL	LLKRYIKKGRKIGVY	WPMGKELRLDG		
	:	: : : :		:     :		
a157	MRNEEKHALRRELRRARAQM	GHQGRLAAGQ	TINRLLKRYIKRGRKIGVY	WPMGKELRLDG		
	10	20	30	40	50	60
m157.pep	FVRAAQKRGAE	LYLPYIEPRSR	RMWFTTYPADGVKQER	KRGRAKLHV	PQFAGRKKRVHDL	
	:			: : :		
a157	FVRAAQKRGAKLYLPYIEPR	SRMWFTTYPESG	MERERIRGRAKLNVPQF	AGRKKIRVHGL		
	70	80	90	100	110	120
m157.pep	NLLLVVGVMDRLGYRLGQAGGY	DATLSAMKYRLQAKTV	GVGFACQLVDR	LPVEAHDRS		
	: : : : :					
a157	SVLLVPLVGI	DREGYRLGQAGGY	DATLAAMKYRLQAKTV	GVGFACQFVDR	LPREPHDLL	
	130	140	150	160	170	180
m157.pep	LDGFVSEAGILCFX					
a157	LDGFVSEAGILCFX					
	190					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 655>:

**g158.seq**

```

1 ATGAAAACCA ATTCAAGA ACTGACCGTA TTTGTTCAAG TGGTGGAAAG
51 CGGCAGCTTC AGCCGTGCGG CGgagcAGTT GGAGAtggCA AATTCTGCCG
101 TAAGCCGCAT CGTCAAACGG CTGGAGGAAA AGTTGGGCGT GAAcCTGctc
151 aACCGCACCA CGCGGCAACT CAATCTGACG GAAGAAGGCG CGCAATATTT
201 CCGCCGCGCG CAGAGAATCC TGCAAGAAAT GGCAGCGGCG GAAACCGAAA
251 TGCTGGCAGT GCACGAAGTA CCGCAAGGCG TGTGCGCGT GGATTCCGCG
301 ATGCcgatgg TGCTGCATCT GCTGGCGCGG CTGGCAGCAA AATTCAACGA
351 ACGCTATCCG CATATCcgac TTTGCTCGT TTCTTCCGAa ggctatatca
401 atctGattGA Acgcaaagtc gAtatTGCCT TACGGGCGCG AGAATTGGAC
451 GATTCCGGGC TGCGTGACG CCATCTGTTT GACAGCCACT TCCGCGtagt
501 cgCCAGTCCT GAATATTTAG CAAAACACGG CACGCCACAA TCTGCAGAAG
551 atcTTGCCAA CCATCAATGT TTAGGCTTCA CAGAACCGG TTCTCTAAAT
601 ACATGGGCGG TTTTAGAtgC GCAGGGAAAT CCTATAAAA TTTACCGCA
651 CTTTACCGCC AGCAGCGGTG AAATCTTACG CTCGTTGTGC CTTTCAAGtt
701 gCGGTATTGC TTGCTTATCA GATTTTGG TTGACAACGA CATCACTGAA
751 GGAAAGTTAA TTCCctatt cgCCGAACAA ACCTCCAATA AAACACACCC

```

```

a158.seq
1  ATGAAACCCA  ATTCAGAAGA  ACTGACCGTA  TTTGTTCAAG  TGGTGGAAAG
51  CGGCAGCTTC  AGCCGTGCGG  CGGAGCAGTT  GGCATGGGTA  AATCTCGCCG
101 TAAGCCGCAT  GCCTCAACCG  CTGGAGGAAA  AGTTGGGTGT  GAACCTGTCT
151 AACCGCACCA  CGCGGCAACT  CAGTCTGACG  GAAGAAGGCG  CGCAATATTT
201 CGCGCGCGCG  CAGAGAAATC  TGCAGAATAA  GGCAGCGGCG  GAAACCGAAA
251 TGCTGGCAGT  GCACGAAATA  CGCAAGGCG  TGTTGCGCGT  GGATTCCGCG
301 ATGCCGATGG  TGCTGCATCT  CCGTGGCGCG  CTGGCAGCAA  AATTCAACGA
351 ACGCTATCCG  CATATCCGAC  TTTGCTCGT  TTCTTCCGAA  GGCTATATCA
401 ATCTGATTGA  ACGCAAAGTC  GATATTGCCT  TACGGGCGCG  AGAATTGGAC
451 GATTCCGGCG  TGCGTGACG  CCATCTGTTT  SACACCCGCT  TCCCGCTAAT
501 CGCGAGTCT  GAATACCTGG  CAAACACCG  CACGCCGCAA  TCTACAGAAG
551 AGCTTGGCGG  CCACCAATGT  TTAGGCTTCA  CCGAACCCGG  TTCTCTAAAT
601 ACATGGGCGG  TTTTAGATGT  CGAGGGAAAT  CCCTATAAGA  TTTCCACCGA
651 CTTTACCGCC  AGCAGCGGTG  AAATCTTACG  CTCGTTGTGC  CTTTCAGGTT
701 GCGGTATTGC  TTGCTTATCA  GATTTTTTGG  TTGACAACGA  CATCGCTGAA
751 GGAAGGTAA  TTCCCTCTGT  CGCGGAACAA  ACCTCCAATA  AAACGCACCC
801 CTTTAATGCT  GTTTATTACA  GCGATAAAGC  CGTCAACCTC  CGCTTACGCG
851 TATTTTTTGA  TTTTTTAGTG  GAGGAACCTG  GAACAATFTC  CTGTGGATAA

```

**a158.pep**

1	MKTNSEELTV	FVQVVESSGF	SRAAEQLMA	NSAVSRIVKR	LEELKGVNLL
51	NRTTRQLSLT	EEGAQYFRR	QRILQEMAAA	ETEMLAVHEI	PGQVLRVDSA
101	MPMVLHLPL	LAAKFNERYP	HIRLSLVSE	GYINLIERKV	DIALRAGELD
151	DSGLRAARHLF	DSRFERYVISP	EYLAKHGTPT	STEELAGHCQ	LGFTPEGSIN
201	TWAVLDAQGN	PYKIFSPHTA	SSGEILRSLC	LSGCGIACLS	DFLVDNDIAE
251	GKLIPLLAEG	TSNKTPEFNA	VYSSDKAVNL	RLRVFLDFLV	EELGNLCC*

	10	20	30	40	50	60
m158.pep	MKTNSEELTVFVQVVEGSGFSFRAAEQLAMANS	AVSRIVKRL	EELKGLGN	LN	NRTTRQ	LSLT
a158	MKTNSEELTVFVQVVEGSGFSFRAAEQLAMANS	AVSRIVKRL	EELKGLGN	LN	NRTTRQ	LSLT
	10	20	30	40	50	60
	70	80	90	100	110	120
m158.pep	EEGAQYFRRAQRILQEMAAAE	TEMLAVHEI	PQGVLS	VD	SAMP	VMVLHL
a158	EEGAQYFRRAQRILQEMAAAE	TEMLAVHEI	PQGVLR	VD	SAMP	VMVLHL
	70	80	90	100	110	120

```

501 CGTCCTGCTG GTGCTTATCC TCGCGCCTA TCTCGAACAG GATAAGGATG
551 TCGAACTCTC GGGCGTATTG AAAGGTGGC AGGACAAACG TTTGGGACAT
601 TTAATCCAAA AGGTGATAGA CAAACCGGAA GACGAATGGA ATGTCGACAA
651 AATGGTGGCG GCTGCCAATA TGTGCGGCGC GCAACTGATG CGCCGTTTCA
701 AAAGCCGGGT CGGACTCAGC CCGCAGCCTT TTGTGAACCA TATCCGCCTG
751 CAAAAAGGCG CGTTGCTGCT GA:AAAAAAC CCGGATTCGG TTTGTGCGGT
801 CGCACTGTCT GTAGGCTTTC AGTCGGAAAC GCACCTCGGC AAGGCGTTCA
851 AACGGCAATA TCACGTTTCG CCGGGTCAAT ACCGAAAGA AggCGGCAG
901 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 664; ORF 160>:

```

m160.pep
1 MDILDKLVDF AQLTGSVDVQ CLLGGQWSVR HETLQREGLV HIVTSGSGYL
51 CIDGETSPRP VSTGDIVFFP RGLGHVLSHD GKCAGESLPD MRQHGAFTVK
101 QCGNGQDMSL FCARFRYDTH ADLMNGLPET VFLNIAHPSL QYVVSMLQLE
151 SKKPLTGTVS MVNALSSVLL VLILRAYLEQ DKDVELSGVL KGWQDKRLGH
201 LIQKVIDKPE DEWNVDMVA AANMSRAQLM RRFKSRVGLS PHAFVNHRL
251 QKGALLLKKN PDSVLSVALS VGFQSETHFG KAFKRYHVS PGQYRKEGGQ
301 K*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m160 / g160 93.4% identity in 301 aa overlap

	10	20	30	40	50	60
m160.pep	MDILDKLVDF	AQLTGSVDVQ	CLLGGQWSVR	HETLQREGLV	HIVTSGSGYL	CIDGETSPRP
g160	MDILDKLVDF	AQLTGSVDVQ	CLLGGQWSVR	HETLQREGLV	HIVTSGSGYL	CIDGETSPRP
	10	20	30	40	50	
m160.pep	VSTGDIVFFP	RGLGHVLSHD	GKCAGESLPD	MRQHGAFTVK	QCGNGQDMSL	FCARFRYDTH
g160	VSTGDIVFFP	RGLGHVLSHD	GKYGESLPD	IRQNGTFMVK	QCGNGLDMSL	FCARFRYDTH
	60	70	80	90	100	110
m160.pep	ADLMNGLPET	VFLNIAHPSL	QYVVSMLQLE	SKKPLTGTVS	MVNALSSVLL	VLILRAYLEQ
g160	ADLMNGLPET	VFLNIAHPSL	QYVVSMLQLE	SEKPLTGTVS	VVNALPSVLL	VLILRAYLEQ
	120	130	140	150	160	170
m160.pep	DKDVELSGVL	KGWQDKRLGH	LIQKVIDKPE	DEWNVDMVA	AANMSRAQLM	RRFKSRVGLS
g160	DKDVELSGVL	KGWQDKRLGH	LIQKVIDKPE	DEWNVDMVA	AANMSRAQLM	RRFKSRVGLS
	180	190	200	210	220	230
m160.pep	PHAFVNHRLQ	KGALLLKKNP	PDSVLSVALS	VGFGQSETHF	GKAFKRYHVS	PGQYRKEGGQ
g160	PHAFVNHRLQ	KGALLLKKNP	PDSVLSVALS	VGFGQSETHF	GKAFKRYHVS	PGQYRKEGGQ
	240	250	260	270	280	290
m160.pep	KX					
g160	KX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 665>:

```

a160.seq
1 ATGGACATTC TGGACAAACT GGTGATTTC GCCCAATTGA CGGGCAGTGT
51 GGATGTGACG TGCCTTTTGG GCGGACAATG GTCGGTACGG CATGAAACCT

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 667>:

```

g161.seq
1  ATGGATACCG CAAAAAAGA CATTTTAGGA TCGGGCTGGA TGCTGGTGGC
51  GGCGGCCTGC TTCACCGTTA TGAACGTATT GATTAAAGAG GCATCGGCAG
101 AATTTGCCCT CGGCAGCGGC GAATTGGTCT TTTGGCGCAT GCTGTTTTC
151 ACCGTACGCG TCGGTGCTGC CGCCGTATTG CGGCGCGACA CCTTCCGCAC
201 GCCCCATTGG AAAAACCACT TAAACCGCAG TATGGTCGGG ACGGGGGCGA
251 TGCTGCTGCT GTTTTACGCG GTAACGCATC TGCCTTTGAC AACCGGCGTT
301 ACCCTGAGTT ACACCTCGTC GATTTTTttg GCGGTATTTT CCTTCCTGAT
351 TTTGAAAGAA CGGATTTCGG TTTACACGCA GCGGTGCTG CTCCTTGGTT
401 TTGCCGGCGT GGTATTGCTG CTTAATCCCT CGTTCGCGAG CGGTCAGGAA
451 CCGGCGGCAC TCGCCGGGCT GCGGGGCGGC GCGATGTCCG GCTGGGCGTA
501 TTTGAAAGTG CGGAACTGT CTTTGGCGGG CGAACCCGCG TGGCGCGTCG
551 TGTTTTACCT TTCCGCAACC GCGTGGC^A TGTCTGcgt ttgggcgacg
601 Ctgaccggct ggCACaccT GTCCTTTcca tcggcagttt ATCtgtCGGG
651 CATCGGCGTG tccgcgCtgA TTGCCCAaCT GtcgatgAcg cGCGcctaca
701 aaGTCGGCGA CAAATTACAG GTTGCTTCGC ttctctaTat gaccgtcGTC
751 TTTTCCGCCC TGTCTGCCGC ATTTTTTCTg ggcgaagagc ttttctggCA
801 GGAAATACTC GGTATGTGCA TCATTATcct CAGCGGCATT TTGAGCAGCA
851 TCCGCCCCAT TGCCTTCAAA CAGCGGCTGC AAGCCTCTT CCGCAAAGA
901 TAA

```

This corresponds to the amino acid sequence <SEQ ID 668; ORF 161.ng>:

```

g161.pep
1  MDTAKKDILG SGWMLVAAAC FTVMNVLIKE ASAKFALGSG ELVFWRMFLS
51  TVTLGAAAVL RRDTFRTPHW KNHLNRSMVG TGAMLLLFYA VTHLPLTTGV
101 TLSYTSSIFL AVFSFLILKE RISVYTOAVL LLGFAGVLL LNPSFRSGQE
151 PAALAGLAGG AMSGWAYLKV RELSLAGEPG WRVVFYLSAT GVAMSSVWAT
201 LTGWHTLSFP SAVYLSGIGV SALIAQLSMT RAYKVGDKFT VASLSYMTVV
251 FSALSAAFFL GEELFWQEIL GMCIIILSGI LSSIPIAFK QRLQALFRQR
301 *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 669>:

```

m161.seq
1  ATGGATACCG CAAAAAAGA CATTTTAGGA TCGGGCTGGA TGCTGGTGGC
51  GGCGGCCTGC TTTACCATTA TGAACGTATT GATTAAAGAG GCATCGGCAG
101 AATTTGCCCT CGGCAGCGGC GAATTGGTCT TTTGGCGCAT GCTGTTTTC
151 ACCGTGCGCG TCGGGGCTGC CGCCGTATTG CGTCGGGACA mCTTCCGCAC
201 GCCCCATTGG AAAAACCACT TAAACCGCAG TATGGTCGGG ACGGGGGCGA
251 TGCTGCTGCT GTTTTACGCG GTAACGCATC TGCCTTTGGC CACTGGCGTT
301 ACCCTGAGTT ACACCTCGTC GATTTTTTTG GCGGTATTTT CCTTCCTGAT
351 TTTGAAAGAA CGGATTTCGG TTTACACGCA GCGGTGCTG CTCCTTGGTT
401 TTGCCGGCGT GGTATTGCTG CTTAATCCCT CGTTCGCGAG CGGTCAGGAA
451 ACGGCGGCAC TCGCCGGGCT GCGGGGCGGC GCGATGTCCG GCTGGGCGTA
501 TTTGAAAGTG CGGAACTGT CTTTGGCGGG CGAACCCGCG TGGCGCGTCG
551 TGTTTTACCT TTCCGTGACA GGTGTGGCGA TGTCTGCGGT TTGGGCGACG
601 CTGACCGGCT GGCACACCCCT GTCCTTTCCA TCGGCAGTTT ATCTGTCTGT
651 CATCGGCGTG TCCGCGCTGA TTGCCCAaCT GTCGATGACG CGCGCCTACA
701 AAGTCGGCGA CAAATTACAG GTTGCTTCGC TTTCTATAT GACCGTCGTT
751 TTTTCCGCTC TGTCTGCCGC ATTTTTTCTG GGCGAAGAGC TTTTCTGGCA
801 GGAAATACTC GGTATGTGCA TCATCATCCT CAGCGGTATT TTGAGCAGCA
851 TCCGCCCCAC TGCCTTCAAA CAGCGGCTGC AATCCTGTG CCGCAAAGA
901 TAA

```

This corresponds to the amino acid sequence <SEQ ID 670; ORF 161>:

```

m161.pep
1  MDTAKKDILG SGWMLVAAAC FTIMNVLIKE ASAKFALGSG ELVFWRMFLS
51  TVALGAAAVL RRDxFRTPHW KNHLNRSMVG TGAMLLLFYA VTHLPLATGV
101 TLSYTSSIFL AVFSFLILKE RISVYTOAVL LLGFAGVLL LNPSFRSGQE

```

851 TCCGCCCCAC TGCCTTCAAA CAGCGGCTGC AATCCCTGTT CCGCCAAAGA  
901 TAA

This corresponds to the amino acid sequence <SEQ ID 672; ORF 161.a>:

a161.pep  
1 MDTAKKDILG SGWMLVAAAC FTIMNVLIKE ASAKFALGSG ELVFWRMLFS  
51 TVALGAAAVL RRDTFRTPHW KNHLNRSVMG TGAMLLLFYA VTHLPLATGV  
101 TLSYTSSIFL AVFSFLILKE RISVYTQAVL LLGFAGVVLL LNPSFRSGQE  
151 TAALAGLAGG AMSGWAYLKV RELSLAGEPG WRVVFYLSVT GVAMSSVWAT  
201 LTGWHLSFPP SAVYLSICIGV SALIAQLSMT RAYKVGDKFT VASLSYMTVV  
251 FSALSAFFL AEELFWQEIL GMCIIILSGI LSSIRPTAFK QRLQSLFRQR  
301 \*

m161/a161 99.3% identity in 300 aa overlap

m161.pep	10	20	30	40	50	60
	MDTAKKDILGSGWMLVAAACFTIMNVLIKEASAKFALGSGELVFWRMLFSTVALGAAAVL					
a161	MDTAKKDILGSGWMLVAAACFTIMNVLIKEASAKFALGSGELVFWRMLFSTVALGAAAVL					
	10	20	30	40	50	60
m161.pep	70	80	90	100	110	120
	RRDXFRTPHWKNHLNRSVMGTGAMLLLFYAVTHLPLATGVTLSTYSIFLAVFSFLILKE					
a161	RRDTFRTPHWKNHLNRSVMGTGAMLLLFYAVTHLPLATGVTLSTYSIFLAVFSFLILKE					
	70	80	90	100	110	120
m161.pep	130	140	150	160	170	180
	RISVYTQAVLLLGFGAGVVLLNPSFRSGQETAALAGLAGGAMSGWAYLKVRELSLAGEPG					
a161	RISVYTQAVLLLGFGAGVVLLNPSFRSGQETAALAGLAGGAMSGWAYLKVRELSLAGEPG					
	130	140	150	160	170	180
m161.pep	190	200	210	220	230	240
	WRVVFYLSVTGVAMSSVWATLTGWHLSFPPSAVYLSICIGVSALIAQLSMTRAYKVGDKFT					
a161	WRVVFYLSVTGVAMSSVWATLTGWHLSFPPSAVYLSICIGVSALIAQLSMTRAYKVGDKFT					
	190	200	210	220	230	240
m161.pep	250	260	270	280	290	300
	VASLSYMTVVFSALSAFFLAEELFWQEILGMCIIILSGILSSIRPTAFKQRLQSLFRQR					
a161	VASLSYMTVVFSALSAFFLAEELFWQEILGMCIIILSGILSSIRPTAFKQRLQSLFRQR					
	250	260	270	280	290	300
m161.pep	X					
	!					
a161	X					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 673>:

g163.seq  
1 ATGGTTATTT TGACGACTTT GTTTTTGTG TGTGTTTGG TGGTATTGGT  
51 TTTAACCGTG CCGGATCAGG TGCAGATGTG gctCGACCGG GCAAAGAAG  
101 TCATTTTAC CGAGTTCAGC TGGTTTATG TTTTAACGTT TTCCATTTT  
151 ctgGGTTTtc tgctGATACT CTCGGTCAGC GGTTCGGAA ACATcagGCT  
201 AGGACGGGAT GAAGATGTGC CGGAATTCGG CTTCCTGTGC TGGCTGGCGA  
251 TGCTGTTTGC GGCCGGGATG GCGTGGGCC TGATGTTTT CGCGTGGCA  
301 GAGCCGTGA TGCATTATT TTCGACATT ACGGTCGGCG CGCCGGAACA  
351 CAGGCAGCAG CAGGCATTGC TGCACACGGT GTTCCATTGG GCGGTTACAG  
401 CCTGGTCGGT GTACGGTACG ATTGCATTGG CTTTGGCTTA TTTCCGTTTC

```

651 CGTGCAGGTT TTGATTATCG CCGCCGTCAT GTCCTCGCC GTCGTTTCGG
701 CAATATCCGG CGTGGGGAAG GCGGTGAAGG TGTGAGCGA GTTGAACCTG
751 GGCCTTGCGT TTTTGCTGCT GTTTTTTGTT TTGCGGCGCG GACCCACTGT
801 TTACCTGTTG TCGGCATTCT GCGACAACAT AGGGAACATC CTCGGAATC
851 TGGTGCGCCT CAGTTTTTAA ACTTATGCGT ACGAACGGGA ACACAAGCCG
901 TGGTTTGAAT CTTGGACGGT GCTTTATTGG GCGTGGTGGT GTTCTTGGGC
951 GCCGTTTGTT GGTGTTGTTA TCGCGCGCAT TTCAAAGGGG CGCACCATCC
1001 GCGAGTTTGT CTTGCGGGTT TTGCTCATCC CCGGCCTGTT CGGCGTTTGT
1051 TGGTTTACCG TCTTCGGCAA TACGCGGATT TGGCTGAATG ACGGGGTTGC
1101 GGGGGGAATG CTCGAAAAGA TGACCTCCTC TCCGGAACCG CTGCTTTTAA
1151 AATTCTTTAA TTACCTCCCC CTGCCGAAAT TGACGAGCAT CGTCAGCCTG
1201 CTGGTCATTT CTCTGTTTTT TGTAACTTCT GCCGATTCCG GGATTTATGT
1251 CCTGAACAAT ATTACCTCTC GGGACAAAGG CTTGAGCGCG CCACGGTGGC
1301 AGGCGGTTAT GTGGGGCGTG CTGATGTCCT CCGTTGCCGT TTTGCTGATG
1351 CGCTCGGGCG GACTCGGCAA CCTGCAGTCT ATGACCTGTA TTGTTTCCCT
1401 GCCGTTTGCC CTGCTGATGC TGATAATGTG TTTAGCCTG TGGAAAAGCT
1451 TGAGTGCGGA TAAGAAATAT TTTGAGACCC GGGTTAACCC TACCAGTGTA
1501 TTTTGACGCG GCGGCAAGTG GAAAGAACGG CTGGTGCAGA TAATGAGCCA
1551 GACGCGAGAG CAGGATATTT TAAATTCCT CAAACAGACT GCATCGCCG
1601 CTATGCACGA GTTGCAACGG GAGCTTTCGG AAGAATACGG CTTGAGCGTC
1651 CGGGTCGATA AAATGTTTCA TCGGGACGAG CCGCAATCG AGTTCGTCAT
1701 TCGGAAAGAG ACGATGCGCG ATTTTATGTA CGGGATTAAG TCTGTCGGG
1751 AGGATGTATC CGACCAAGTG ATTAACGACG GCAAGCTGCC GCATATCCG
1801 CATCAGACAA CTTACAAACC CTACGCTTAT TTTTTCGACG GCGCGTCGG
1851 GTACGATGTG CAGTATATGA ACAAGGACGA GCTGATTGCC GACATTTTGA
1901 AAAACTACGA ACGTTATTTG ATGTTGTTGG ATGATGTCGG TCAGGAACCTG
1951 ATGGCGCACG AGCAGGTGGA ATTGGCAGAG TAA

```

This corresponds to the amino acid sequence <SEQ ID 676; ORF 163>:

```

m163.pep
1  MVILTTLFFV CVLVVLVLTVPDQVQMWLDR AKEVIETEF S WFYVLTFSIF
51  LGFLLILSV S SLGNIRLGRD EDVPEFGFLS WLAMLFAGM GVGLMFFGVA
101 EPLMHYFSDI TAGTPEHRQQ QALLHTVFHW GVHAWSVYGT IALALAYFGF
151 RYKLPLALRS CFYPLLKEKI SGRFGDAIDI MALLATFFGI ITTLGFGASQ
201 LGAGLQEMGW IAENSFVSQV LIIAAVMSLA VVSAISGVGK GVKVLSLNL
251 GLAFLLLFFV LAAGPTVYLL SAFGDNIGNY LGNLVRLSFK TYAYEREHKP
301 WFESWTVLYW AWWCSWAPFV GLFIARISKG RTIREFVFGV LLIPGLFGVL
351 WFTVFGNTAI WLNDGVAGGM LEKMTSSPET LLFKFFNYLP LPELTSIVSL
401 LVISLFFVTS ADSGIYVLNN ITS RDKGLSA PRWQAVMWGV LMSAVAVLLM
451 RSGGLGNLQS MTLIVSLPFA LLMLIMCFSL WKGLSADKKY FETRVNPTSV
501 FWTGGKWKER LVQIMSQTQE QDILKFLKQT ASPAMHELQR ELSEEYGLSV
551 RVDKMFHRDE PAIEFVIRKE TMRDFMYGIK SVGDVSDQL INDGKLPHIR
601 HQTTPYKPYAY FFDGRVGYDV QYMNKDELIA DILKNYERYL MLLDDVQGEL
651 MAHEQVELAE *

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m163 / g163 98.6% identity in 660 aa overlap

	10	20	30	40	50	60
m163.pep	MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIETEF S WFYVLTFSIFLGFLILSVS					
g163	MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIETEF S WFYVLTFSIFLGFLILSVS					
	10	20	30	40	50	60
m163.pep	SLGNIRLGRDEEDVPEFGFLSWLAMLFAGMGVGLMFFGVAEPLMHYFSDITAGTPEHRQQ					
g163	GLGNIRLGRDEEDVPEFGFLSWLAMLFAGMGVGLMFFGVAEPLMHYFSDITVGAPEHRQQ					
	70	80	90	100	110	120
m163.pep	SLGNIRLGRDEEDVPEFGFLSWLAMLFAGMGVGLMFFGVAEPLMHYFSDITAGTPEHRQQ					
g163	GLGNIRLGRDEEDVPEFGFLSWLAMLFAGMGVGLMFFGVAEPLMHYFSDITVGAPEHRQQ					
	70	80	90	100	110	120
	130	140	150	160	170	180



```

201 CGGACGGGAT GAAGATGTGC CGGAATTCGG CTTCTGTCTG TGGCTGGCGA
251 TGCTGTTTGC GGCCGGGATG GGCGTGGGTC TGATGTTTTT CGGCGTGGCA
301 GAGCCGTTGA TGCATTATTT TTCGGACATT ACGGCCGGCA CGCCGGAACA
351 CAGGCAGCAG CAGGCATTGC TGCACACGGT GTTCCATTGG GGCCTTCACG
401 CTTGGTCGGT GTACGGTACG ATTGCATTGG CTTTGGCTTA TTTCGGTTTC
451 CGCTACAAGC TGCCGCTTGC CCTGCGTTCT TGTTTTTACC CCCTGTTGAA
501 AGAAAAAATT TCCGGAAGGT TCGCCGATGC CATTGATATT ATGGCGTTGC
551 TTGCTACTTT TTTCGGCATC ATCACCACAT TGGGGTTCGG GGCTTCGCAA
601 CTGGGCGCCG GATTGCAGGA AATAGGCTGG ATTGCCGAAA ACAGCTTCAG
651 CGTGCAGGTT TTGATTATCG CCGCCGTCAT GTCCCTCGCC GTCGTTTCGG
701 CAATATCCGG CGTGGGGAAG GGTGTGAAGG TGTGAGCGA GTTGAACCTG
751 GGTCTTGCGT TTTTGTGCT GTTTTTTGT TGGCGGCGG GTCCCACTGT
801 TTACCTGTTG TCGGCATTCT GCGACAACAT AGGGAACAC CTGCGAAATC
851 TGGTGCGCCT CAGTTTTAAA ACTTATGCGT ACGAACGGGA ACACAAGCCG
901 TGGTTTGAAT CTGGACGGT GCTTTATTGG CGGTGGTGGT GTTCTTGGGC
951 GCCGTTTGTG GGTTTGTTTA TCGCGCGCAT TTCAAAGGGG CGCACCATCC
1001 GCGAGTTTGT CTTCGGGTTT TTGCTCATCC CCGGCCTGTT CGGCGTTTGT
1051 TGGTTTACCG TCTTCGGCAA TACGGCGATT TGGCTGAATG ACGGGGTTGC
1101 GGGGGGAGTG CTCGAAAAGA TGACCTCCTC TCCGGAACG CTGCTTTTAA
1151 AATTCTTTAA TTACCTCCCC CTGCCGAAT TGACGAGCAT CGTCAGCCTG
1201 CTGGTCATT CTCTGTTTTT TGAACCTCT GCGGATTCCG GGATTTATGT
1251 CCTGAACAA ATTACCTCTC GGGACAAAGG CTGAGCGCG CCACGGTGGC
1301 AGGCGGTTAT GTGGGGCGTG CTGATGCTG CCGTTGCCGT TTGCTGATG
1351 CGCTCGGGCG GACTCGGCAA CCTGCAGTCT ATGACCCTGA TTGTTTCCCT
1401 GCCGTTTGCC CTGCTGATGC TGATAATGT TTTACGCCTG TGGAAAGGAT
1451 TGAGTGCAGA TAAGAAATAT TTTGAGACCC GGGTTAACCC TACCGTGTGA
1501 TTTTGGACGG GCGGCAAGTG GAAAGAACGG CTGGTGCAGA TAATGAGCCA
1551 GACGCAGGAG CAGGATATTT TAAATTCCT CAAACATACC GCATCGCCCG
1601 CTATGCACGA GTTACAACGG GAGCTTTCGG AAGAATACGG CTGAGCGTC
1651 CGGGTCGATA AGATGTTTCA TCAGGACGAG CCCGCAATCG AGTTCGTCAT
1701 TCGGAAAGAG ACGATGCGCG ATTTTATGTA CGGGATTAAG TCTGTCGGGC
1751 AGGATGTATC CGACCAAGTG ATTAACGACG GCAAGCTGCC GCATATCCGG
1801 CATCAGACAA CTTACAAACC CTACGCTTAT TTTTTCGACG GCGCGGTCGG
1851 GTACGATGTG CAGTATATGA ACAAGGACGA GCTGATTGCC GACATTTTGA
1901 AAACTACGA ACGTTATTTG ATGTTGTTGG ATGATGTCGG TCAGGAACTG
1951 ATGGCGCACG AGCAGGTGGA ATTGGCAGAG TAA

```

This corresponds to the amino acid sequence <SEQ ID 678; ORF 163.a>:

```

a163.pep
1  MVILTTLFFV CVLVVLVLTVPDQVQMWLDR AKEVIFTEFS WFYVLTFSIF
51  LGFLLILSVS SLGNIRLGRD EDVPEFGFLS WLAMLFAGM GVGLMFFGVA
101 EPLMHYFSDI TAGTPEHRQQ QALLHTVFHW GVHAWSVYGT IALALAYFGF
151 RYKLPLALRS CFYPLLKEKI SGRFGDAIDI MALLATFFGI ITTLGFGASQ
201 LGAGLQEIOW IAENSFSVQV LIIAAVMSLA VVSAISGVGR GVKVLSELNL
251 GLAFLLLFFV LAAGPTVYLL SAFGDNIGNY LGNLVRLSEK TYAYEREHKP
301 WFSWTVLYW AWWCSWAPFV GLFIARISKG RTIREFVFGV LLIPGLFGLV
351 WFTVFGNTAI WLNDGVAGGV LEKMTSSPET LLFKFFNYLP LPELTSIVSL
401 LVISLFFVTS ADSGIYVLNN ITS RDKGLSA PRWQAVMWGV LMSAVAVLLM
451 RSGGLGNLQS MTLIVSLPFA LLMLIMCFSL WKGLSADKKY FETRVNPTSV
501 FWTGGKWKER LVQIMSQTQE QDILKFLKHT ASPAMHELQR ELSEEYGLSV
551 RVDKMFHQDE PAIEFVIRKE TMRDFMYGIK SVQDQVSDQL INDGKLPHIR
601 HQTYYKPYAY FFDGRVGYDV QYMNKDELIA DILKNYERYL MLLDDVGOEL
651 MAHEQVELAE *

```

m163/a163 99.4% identity in 660 aa overlap

```

              10      20      30      40      50      60
m163.pep      MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIFTEFSWFYVLTFSIFLGFLILSVS
              |||
a163           MVILTTLFFVCVLVVLVLTVPDQVQMWLDRAKEVIFTEFSWFYVLTFSIFLGFLILSVS
              10      20      30      40      50      60

              70      80      90     100     110     120
m163.pep      SLGNIRLGRDEDVPEFGFLSWLAMLFAGMGVGLMFFGVAEPLMHYFSDITAGTPEHRQQ

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 679>:

```

g164.seq (partial)
1  ..ATGAACACAT TTTTGAAAAA CAGCGAATAC GCGTATATCC TGAACGACTG
51  CAAGGCGCGC TTCCTGTTTCG CCTCGGCCCG CCTGTCAAAA GAATTGGCCG
101 GCCTGAAGGC GCAAACGCCG GTCGAAAAAA TCATTTGGAC GGACAAAAGC
151 CCGCCGCGCG GCGAAACGGC GGAAGGCGAT GCCTTTTGTG AAAACGTGCG
201 CCGCTTCCCG GAAAACCCG ACTTGGGCCG CCAACCCCGG ATAAATGATT
251 TGGCACACAT CATCTACACC TCCGGCACGA CGGGGCATCC CAAAGCGCGC
301 CTATCAGTT ACGCCAACCT GTTCGCCAAC CTGAACGGCA TCGAACGCAT
351 CTTtaaAATT TCCAAACGCG ACCGCTTTAT CGTTTTcctg ccgatgTTCC
401 ACAGCTTCAC GCTGACGGCT ATGGTGCTGC TGGCGATTTA TATGGCGTGT
451 TCGATTATTT TGGTCAAAtc cgttttCCcC ttttccaaCG TTTTGAAACA
501 GGCCCTGCTC AAACGCGCAA CCGTGTTTTT GGGCGTACCC GCGATTTCAC
551 CCGCGATGAG CAAGGCAAAA ATCCCTTGGT ATTTCAGATG GTTCAACCGC
601 ATCCGCTGTG TTATCAGCGG CGGCGCGCCT TTGGCGGAAC AAACCATCCT
651 CGATTTTAAA GCCAAGTTCC CCGCGCCCAA ATTGCTGGAA GGCTACGGAC
701 TGAGCGAAGC CTCGCCCGTC GTCGCCGTCA TACGCCCGA ACGGCAAAAA
751 GCCGCGAGCG TCGGCATCCC CCTGCCCGGT TTGGAAGCCA AAGCCGTCGA
801 TGAAGAAATTG GTCGAAGTGC CGCGCGGCGA AGTGGGCGAA CTGATCGTCA
851 GGGGCGGTTT GGTGATGCGG GGCTACCTCA ATATGCCTGC CGCCACCGAT
901 GAAACCATCG TCAACGGCTG GTTGAAAACG GCGGATTTCC TTACCATAGA
951 CGAGGACGGC TTTATCTTTA TCGTCGACCG CAAAAAAGAT TTGATTATTT
1001 CCAAAGGTCA AAACGTCTAT CCGCGCGAGA TCGAAGAAGA AATCCACAAA
1051 CTCGATGCCG TCGAAGCCGC CGCCGTCATC GGCGTGAAAG ACCGTTATGC
1101 CGACGAGGAA ATCGTCGCCT TCGTCCAATT GAAGGAAGGT ATGGATTGG
1151 GCGAGGACGA aatccgccc caccTGCGTA CCGTGTGGC AAATTTCAAA
1201 ATCCCCAAAC AGATCCACTT TAAAGACGGG CTGCCGCGCA ACCTACGGG
1251 CAAAGTATTG AAACGGGTGC TGAAGGAGCA GTTTGAAGGA AACAAATGA

```

This corresponds to the amino acid sequence <SEQ ID 680; ORF 164.ng>:

```

g164.pep (partial)
1  ..MNTFLKNSEY AYILNDCKAR FLFASAGLSK ELAGLKAQTP VEKIIWTDKS
51  RPAGETAEGD AFFENVRRFP EKPDLGRQPR INDLAHIIYT SGTTHPKGA
101 LISYANLFAN LNGIERIFKI SKRDRFIVFL PMEHSFTLTA MVLPLIYMAC
151 SIILVKSVPF FSNVLKQALL KRATVFLGVP AIYTAMSKAK IPWYFRWFNR
201 IRLFISGGAP LAEQTILDFK AKFPRAKLE GYGLSEASPV VAVNTPERQK
251 ARSVGIPLPG LEAKAVDEEL VEVPRGEVGE LIVRGGSVMR GYLNMPAATD
301 ETIVNGWLT GDFVTIDEDG FIFIVDRKRD LIISKQNVY PREIEEIIHK
351 LDAVEAAAVI GVKDRYADEE IVAFVQLKEG MDLGEDEIRR HLRTVLANFK
401 IPKQIHFKDG LPRNATGKVL KRVLKEQFEG NK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 681>:

```

m164.seq
1  ATGAACCGGA CTTATGCCAA TTTCTACGAA ATGCTCGCCG CCGCCTGCCG
51  CAAAAACGGA AACGGCACGG CAGTGTTCGA CGGCAAGGAA AAAACCGCCT
101 ACCGCGCGCT CAAGCAGGAG GCCGAAGCCG TCGCGCGGTA TCTGCAAAAT
151 ATCGGCGTGA AGTTCGCGCA CACGGTCGCG CTGGCGGTTT CCAATTCCAC
201 AGAATTTATT ACCGCTATT TCGCCATCTC CGCCATCGGC GCGGTCGCGC
251 TACCGATGAA CACATTTTGT AAAAACAGCG AATACGCGTA TATCCTGAAC
301 GACTGCAAGG CGCGCTTCCT GTTCGCCTCG GCCGCGCTGT CAAAAGAATT
351 GGCGGGCTTG AAGGCGCAAA CGCCCGTCGA AAAAATCATT TGGACGGACA
401 AAAGCCGTCG GACCGGCGAA ACGCGGAAG GCGATGCCTT TTTGAAGAC
451 GTGCGCGGCT TCCCCGAAAA ACCCGACTTG GGCGGCCAAC CCCGGATAAA
501 TGATTTGGCA CACATCATCT ACACCTCCGG CACGACGGG CATCCCAAAG
551 GCGCGCTAAT CAGTTACGCC AACCTGTTTC CCAACCTGAA CGGCATCGAA
601 CGCATCTTTA AAATTTCCAA GCGCGACCGC TTTATCGTTT TCCTGCCGAT
651 GTTCCACAGC TTCACGCTGA CGGCTATGGT GCTGCTGCCG ATTTATATGG
701 CGTGTTCGAT TATTTTGGTC AAATCCGTTT TTCCGTTTTC CAACGTTTTC
751 AAACAGACAC TGCTCAAACG CGCGACCGTG TTTTGGGCG TACCCGCGAT
801 TTACACCGCG ATGAGCAAGG CGAAAATCCC TTGGTATTTC AGATGGTTCA
851 ACCGATTTCG CTTGTTTATC AGCGGCGGCG CGCCTTTGGC GGAACAAACC
901 ATCTCGATT TCAAAGCCAA GTTCCCCCGC GCCAAATTGC TGAAGGCTA

```

467

```

m164.pep  VEVPRGEVGELIVRGGSVMRGYLNMPAATDEIVNGWLKTGDFVTIDEDGFIFIVDRKKD
          |||
g164      VEVPRGEVGELIVRGGSVMRGYLNMPAATDEIVNGWLKTGDFVTIDEDGFIFIVDRKKD
          280      290      300      310      320      330

          420      430      440      450      460      470
m164.pep  LIISKQNVYPREIEEEIYKLDAAVEAAVIGVKDRYADEEIVAFVQLKEGMDLGENEIRR
          |||
g164      LIISKQNVYPREIEEEIHKLDAAVEAAVIGVKDRYADEEIVAFVQLKEGMDLGEDEIRR
          340      350      360      370      380      390

          480      490      500      510
m164.pep  HLRTVLANFKIPKQIHFKDGLPRNATGKVLKRVLKEQFDGNKX
          |||
g164      HLRTVLANFKIPKQIHFKDGLPRNATGKVLKRVLKEQFEGNKX
          400      410      420      430

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 683>:

```

a164.seq
1  ATGAACCGGA CTTATGCCAA TTTCTACGAA ATGCTGACCG CCGCCTGCCG
51  CAAAAACGGC AACGGCACGG CAGTGTTCGA CGGCAAGGAA AAAACCGCCT
101 ACCGCGCGCT CAAGCAGGAA GCCGAAGCCG TTGCGGCGTA TCTGCAAAAT
151 ATCGGCGTGA AGTTCGGCGA CACGGTCGCG CTGGCGGTTT CCAATTCAC
201 GGAATTATT ACCGCTTATT TCGCCGTATC CGCCATCGGC GCGGTTGCCG
251 TACCGATGAA CACATTTTGT AAAACAGCG AATACGCGTA TATCTGAAC
301 GACTGCAAGG CGCGCTTCCT GTTCGCTCG GCCGGCCTGT CAAAAGAATT
351 GCGGGGCTTG AAGGCGCAA CGCCGTCGA AAAATCATT TGGACGGGCC
401 AAAGCCGTCC GGACGGCGAA ATGGCGGAAG GCGATGCCTT TTTGAAGAC
451 GTGCGCGCT TCCCGAAAA ACCCGACTTG GGCGGCAAC CCCGGATAAA
501 TGATTGGCA CACATCATCT ACACCTCCGG CACGACGGGG CATCCCAAAG
551 GTGCGTAAT CAGCTACGCC AACCTGTTCT CCAACCTGAA CGGCATCGAA
601 CGCATCTTTA AAATCTCCAA GCGCGACCGC TTTATCGTTT TCCTGCCGAT
651 GTTCCACAGC TTCACGCTGA CCGCTAAGG GCTGCTGCCG ATTTATATGG
701 CGTGTTCGAT TATTTTGGTC AAATCCGTTT TCCCTTTTC CAACGTTTGT
751 AAACAGGCAC TGCTCAAACG CGCGACCGTG TTTTGGGCG TGCCCGCGAT
801 TTACACCGCG ATGAGCAAGA CGAAATCCC TTGGTATTTC AGATGGTTCA
851 ACCGCATCCG CCTGTTTATC AGCGCGGAG CACCTTTGGC GGAACAAACC
901 ATCCTCGATT TCAAAGCCAA GTTCCCCCGC GCCAAATTGC TGGAAGGCTA
951 CGGACTGAGC GAAGCCTCGC CCGTCGTCGC CGTCAATACG CCCGAGAGGC
1001 AAAAAGCCCG CAGCGTCGGC ATCCCCCTGC CCGGTTTGGG AGTCAAAGCC
1051 GTCGATGAAG AATGGTTCGA AGTGCCGCGC GGCGAAGTGG GCGAACTGAT
1101 CGTCAGGGGC GGTTCGGTGA TGCGGGGCTA CCTCAATATG CCTGCCGCCA
1151 CCGATGAAAC CATCGTCAAC GGCTGGTTGA AAACGGGCGA TTTCTGTACC
1201 ATAGACGAAG ACGGCTTTAT CTTTATCGTC GACCGCAAAA AAGATTTGAT
1251 TATTTCCAAA GGTCAAAATG TCTATCCGCG CGAAATCGAA GAAGAAATCT
1301 ACAAATCGA TGCCGTCGAA GCCGCCGCG TCATCGGCGT GAAAGACCGT
1351 TATGCCGACG AGGAAATCGT CGCCTTCGTC CAATTGAAGG AAGGTATGGA
1401 TTTGGGCGAG AACGAAATCC GCCGCCACCT CGGTACCGTG CTGGCAAAAT
1451 TCAAAATCCC CAAACAAATC CACTTTAAAG ACGGGCTGCC GCGCAACGCT
1501 ACGGGCAAGG TATTGAAACG GGTGTTGAAG GAGCAGTTTG ACGGAAACAA
1551 ATGA

```

This corresponds to the amino acid sequence <SEQ ID 684; ORF 164.a>:

```

a164.pep
1  MNRTYANFYE MLTAACRKNG NGTAVFDGKE KTAYRALKQE AEAVAAAYLQN
51  IGVKFGDTVA LAVSNSTEFI TAYFAVSAIG AVAVPMNTFL KNSEYAYILN
101 DCKARFLFAS AGLSKELAGL KAQTPVEKII WTGQSRPDGE MAEGDAFFED
151 VRRFPEKPD LGRQPRINDLA HIIYTS GTTG HPGALISYA NLFANLNGIE
201 RIFKISKRDR FIVFLPMFHS FTLTAMVLLP IYMACSILV KSVFPFSNVL
251 KQALLKRATV FLGVPPIYTA MSKTIPWYF RWFENRIRLFI SGGAPLAEQT
301 ILDFKAKFPR AKLLEGYGLS EASPVVAVNT PERQKARSVG IPLPGLEVKA
351 VDEELVEVPR GEVGELIVRG GSVMRGYLNM PAATDEIVN GWLKTGDFVT
401 IDEDGFIFIV DRKKDLISK QNVYPREIE EEIYKLDAAVE AAAVIGVKDR
451 YADEEIVAFV QLKEGMDLGE NEIRHLRTV LANFKIPKQI HFKDGLPRNA

```

```

51  GACTTTGGGC GTTTTGCTCA AAGAACTCGA ACCGTCTTGG GAAATCACCC
101 TGATTGAACG CTTGGAagat gTGGCGTTGG AATCGTCAAA cGCGTGGAAc
151 AACGcCGgca CGGGGCATTC CGcGCTGTGc GAATTGAACT AtgcgccGCT
201 GGGtgccggac ggcgtcatCA ATCCGGCGCg cgCCCTGAAT ATTGCCGAAC
251 AGTTTCATGT CAGCCGCCAG TTTTGGGcga cgcgtGTCGc GGAAGGCAAG
301 TTGGAagaCA ATTCCTTCAT CAATGCcgtg ccgcatatGT Ctttggtgat
351 gAacgaagac cactgCCgtt acCTGCAAAA ACGCTATGAT GTGTTTAAAA
401 CGCAGAAACT TTTTGAAAAT ATGGAATTTT CCACCGATCG GAACAAAAAT
451 TCCGATTGGG CtcgcCTGAT TATGCGCGGC CGGgacgaaa ACCAACCCGT
501 CGCCGCCAAC TATTCCGCCG Aaggcacgga tgtcgATTTC GGACGGCTGA
551 CGCGCCAGAT GGTGAAATAT TTGCAGGGCA AGGGCGTAAA AACCGAGTTC
601 AACCGCCACG TCGAAGACAT CAAACGCGAA TCCGACGGCG CGTGGGTGCT
651 CAAAACCGCC GATACCCGCA ACCCAGACTG GCAGCTCACC CTCCGCACCC
701 GCTTCCTCTT CCTCGGCGCG GCGGCGCGCG CACTGACCCCT GCTGCAAAAA
751 TCCGGCATCC CCGAAGGCAA AGGCTACGGC GGCTTACCCG TGTCGGGCCT
801 GTTCTTCCGC AACAGCAACC CCGAAACCGC CGAACAACAC AACGCCAAAG
851 TGTACGGGCA GGCTTCCGTC GCGCGCGCGC CGATGTCCGT CCCGCACCTC
901 GACACACGCA ACGTAGACGG CAAACGCAC CTTATGTTTC GTCCCTACGC
951 AGGTTTCCGT TCCAACCTCC TCAAGCAAGG CTCGTTTATG GATTTGCCCG
1001 TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGCGCGC CGGCTGGGCG
1051 AATATGCCCG TGACCAAATA CcTGCTGGGc gAaTTGCgtt aa

```

This corresponds to the amino acid sequence <SEQ ID 686; ORF 165.ng>:

```

g165.pep
1  MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
51  NAGTGHSALC ELNYAPLGAD GVINPARALN IAEQFHVSRO FWATLVAEGK
101 LEDNSFINAV PHMSLVMNED HCRYLQKRYD VFKTQKLFEN MEFSTDRNKI
151 SDWAPLIMRG RDNQPVAAAN YSAEGTDVDF GRLTRQMVKY LQKGKVKTEF
201 NRHVEDIKRE SDGAWVLKTA DTRNPDWQLT LRTRELFLGA GGGALTLLQK
251 SGIPEGKGYG GLPVSGLFFR NSNPETAEOH NAKVYQASV GAPPMSVPHL
301 DTRNVDGKRH LMFPGYAGFR SNFLKQGSFM DLPLSIHMDN LYPMLRAGWA
351 NMPLTKYLLG ELR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 687>:

```

m165.seq (partial)
1  ATGGCTGAAG CGACAGACGT TGTCTGGTG GCGGCGGCA TTATGAGCGC
51  GACTTTGGGC GTTTTGCTCA AAGAACTCGA ACCGTCTTGG GAAATCACCC
101 TGATTGAACG CTTGGAAGAT GTGGCGTTGG AATCGTCAAA CGCGTGGAAc
151 AACGCCGCA CGGGGCATTC CGCGCTGTGc GAATTGAACT ATGCGCCGTT
201 GGGTGCAAAAT GGGATTATCG ATCCGGCGCG CGCCCTCAAT ATTGCCGAAC
251 AGTTTCATGT CAGCCGCCAG TTTTGGGCGA CGCTGGTCCG GGAAGGCAAG
301 TTGGAAGACA ATTCCTTCAT CAATGCCGTG CCGCATATGT CTTTGGTGAT
351 GAATGAAGAC CATTGTTCTT ATCTTCAAAA ACGTTATGAC GCGTTTAAAA
401 CCCAAAAACT TTTTGAAAAT ATGGAATTTT CCACCGATCG GAACAAAAAT
451 TCCGATTGGG CTCCGCTGAT GATGCGCGGC CGGGACGAAA ACCAACCCGT
501 CGCCGCCAAC TACTCCGCCG AAGgTACGGA TGTCGATTTC GGACGGCTGA
551 CGCGCCAAAT GGTGAAATAT TTGCAGGGCA AGGGCGTAAA AACCGAGTTC
601 AACCGCCACG TCGAAGACAT CAAACGCGAA TCCGACGGCG CGTGGGTGCT
651 CAAAACCGCC GATACCCGCA ACCCGACGGC GCAGCTCACC CTCCGTACCC
701 GCTTCCTCTT CCTCGGCGCG GCGGCGCGCG CGCTGACCCCT GCTGCAAAAA
751 TCCGGCATCC CCGAAGGCAA AGGCTACGGC GGCTTCCCGG TGTCGGGCCT
801 GTTCTTCCGC AACAGCAACC CCGAAACCGC CGAACAACAC AACGCCAAAG
851 TGTACGGGCA GGCTTCCGTC GCGCGCGCGC CGATGTCCGT CCCGCACCTC
901 GACACACGCA ACGTGGACGG CAAACGCCAC CTTATGTTTC GCCCTTACGC
951 AGGCTTCCGT TCCAACCTCC TCAAGCAAGG CTCGTTTATG GATTTGCCCG
1001 TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGTGCGC CGGCTGGGCG
1051 AATATGCCCG TGACCAA...

```

This corresponds to the amino acid sequence <SEQ ID 688; ORF 165>:

```

m165.pep (partial)
1  MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
51  NAGTGHSALC ELNYAPLGAN GIIDPARALN IAEQFHVSRO FWATLVAEGK
101 LEDNSFINAV PHMSLVMNED HCRYLQKRYD VFKTQKLFEN MEFSTDRNKI
151 SDWAPLIMRG RDNQPVAAAN YSAEGTDVDF GRLTRQMVKY LQKGKVKTEF

```

```

601 AACCGCCACG TCGAAGACAT CAAACGCGAA TCCGACGGCG CGTGGGTGCT
651 CAAAACCGCC GATACCGGCA ACCCGGACGG GCAGCTCACC CTCCGTACCC
701 GCTTCTCTT CCTCGGCGCG GCGGCGGCG CGCTGACCCT GCTGCAAAAA
751 TCCGGCATCC CCGAAGGCAA AGGCTACGGT GGCTTTCCCG TGTCGGGCT
801 GTTCTTCCGC AACAGCAACC CCGAAACCGC CGAACAACAC AACGCCAAG
851 TGTACGGGCA GGCTTCCGTC GCGCGCGCCG CGATGTCCGT CCGCACCTC
901 GACACACGCA ACGTGGACGG CAAACGCCAC CTTATGTTTC GCCCTTACGC
951 AGGCTTCCGT TCCAACCTTC TCAAGCAAGG CTCACTTATG GATTTGCCGC
1001 TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGCGCGC CGGCTGGGCG
1051 AATATGCCGC TGACCAAATA CCTGCTGGGC GAATTGCGTA AAACCAAGA
1101 AGAACGCTTC GCCTCCCTGC TGAATACTA CCCCAGGCA AACCCGACG
1151 ACTGGGAAC CATCACCGCA GGGCAACGCG TTCAATCAT TAAAAAGAC
1201 TCCGAAAAAG GCGGCGTGTT GCAGTTGGT ACGGAGATTG TCGCACACGC
1251 CGACGGCTCG CTCGCCGCAT TGCTGGGCGC GTCGCCGGGC GCATCGACCC
1301 CCGTGCCGCT GATGATCCGG CTGATGCACC AATGCTTCCC CGAACGCACC
1351 CCGTCTTGGG AAGGCCGTCT GAAAGAGCTG GTACCGGGTT ACGGCATCAA
1401 GTTGAACGAA AACCCCGAAA GGGCGGATGA AATTATCGCC TATACCGCGA
1451 AAGTGTGGA TATTTAA

```

This corresponds to the amino acid sequence <SEQ ID 690; ORF 165.a>:

```

a165.pep
1 MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
51 NAGTGHSA LC ELNYAPLGAN GIIDPARALN IAEQFHVSRQ FWATLVAEGK
101 LEDNSFINAV PHMSLVMNED HCSYLQKRYD AFKTQKLFEN MEFSTDRNKI
151 SDWAPLMMRG RDNQPVAAAN YSAEGTDVDF GRLTRQMVKY LQKGKVKTEF
201 NRHVEDIKRE SDGAWVLKTA DTRNPDGQLT LTRFLFLGA GGGALTLLQK
251 SGIEGKGYG GFPVSGLFFR NSNPETAQH NAKVYQASV GAPPMSVPHL
301 LTRNVGDKRH LMFPGYAGFR SNFLKQGS LMPLSIHMDN LYPMLRAGWA
351 NMPLTKYLLG ELRKTEERF ASLLEYYPEA NPDDWELITA GQRVQI IKKD
401 SEKGGVLQFG TEIVAHADGS LAALLGASPG ASTAVPLMIR LMHQCFPERT
451 PSWEGRLKEL VPGYGIKLE NPERADEIIA YTAKVLDI*

```

m165/a165 99.7% identity in 356 aa overlap

m165.pep	10	20	30	40	50	60
a165	10	20	30	40	50	60
m165.pep	70	80	90	100	110	120
a165	70	80	90	100	110	120
m165.pep	130	140	150	160	170	180
a165	130	140	150	160	170	180
m165.pep	190	200	210	220	230	240
a165	190	200	210	220	230	240
m165.pep	250	260	270	280	290	300
a165	250	260	270	280	290	300

473

```

851 TGTACGGGCA GGCTTCGGTC GCGCGCGCGC CGATGTCCGT CCCGCACCTC
901 GACACACGCA ACGTGGACGG CAAACGCCAC CTTATGTTCC GCCCTTACGC
951 AGGCTTCCGT TCCAACCTCC TCAAGCAAGG CTCGCTTATG GATTTGCCGC
1001 TGTCCATCCA TATGGACAAC CTCTATCCTA TGCTGTGCGC CGGCTGGGCG
1051 AATATGCCGC TGACCAAATA CCTGCTGGGC GAATTGGCGTA AAACCAAAGA
1101 AGAACGCTTC GCCTCCCTGC TGGAAATACTA CCCCAGGGCA AACCCGACG
1151 ACTGGGAACT CATCACCGCA GGGCAACGCG TCCAAATCAT TAAAAAGAC
1201 TCCGAAAAG GCGGCGTGCT CCAGTTTGGT ACGGAGATTG TCGCCACGC
1251 CGACGGCTCA CTCGCCGAT TGCTGGGCGC GTCGCCGGG GCATCGACCG
1301 CTGTGCCGCT GATGATCCGG CTGATGCACC AATGCTTCCC CGAGCGCGCC
1351 CCGTCTTGGG AAGACCGTCT GAAAGAGCTG GTACCGGGT ACGGCATCAA
1401 GTTGAACGAA AACCTGAAA GGGCGGATGA AATTATCGCC TATACCGCGA
1451 AAGTATTGGA TATTTAA

```

This corresponds to the amino acid sequence <SEQ ID 694; ORF 165-1>:

**m165-1.pep**

```

1 MAEATDVVLV GGGIMSATLG VLLKELEPSW EITLIERLED VALESSNAWN
51 NAGTGHSALC ELNYAPLGAN GIIDPARALN IAEQFHVSQ FWATLVAEGK
101 LEDNSFINAV PHMSLVMNED HCSYLQKRYD AFKTQKLFEN MEFSTDRNKI
151 SDWAPLMMRG RDENQPVAA NYSAEQTDVDF GRLTRQMVKY LQKGKVKTEF
201 NRHVEDIKRE SDGAWVLKTA DTRNPDGQLT LTRFLFLGA GGGALTLLQK
251 SGIPEGKGYG GPFVSGLFER NSNPETAQH NAKVYGQASV GAPPMSVPHL
301 DTRNVDGKRH LMFGPYAGFR SNFLKQGSIM DLPLSIHMDN LYPMLCAGWA
351 NMPLTKYLLG ELRKTKEERF ASLLEYYPEA NPDDWELITA GQRVQIHKD
401 SEKGGVLQFG TEIVAHADGS LAALLGASPG ASTAVPLMIR LMPVTPERA
451 PSWEDRLKEL VPGYGIKLE NPERADEIIA YTAKVLDI*

```

**m165-1/g165-1** 89.7% identity in 428 aa overlap

	10	20	30	40	50	60
m165-1.pep	MAEATDVVLVGGGIMSATLG	VLLKELEPSWEITLIERLED	VALESSNAWN	NAGTGHSALC		
g165-1	MAEATDVVLVGGGIMSATLG	VLLKELEPSWEITLIERLED	VALESSNAWN	NAGTGHSALC		
	10	20	30	40	50	60
	70	80	90	100	110	120
m165-1.pep	ELNYAPLGANGIIDPARALN	IAEQFHVSQFWATLVAEGK	LEDNSFINAVPHMSLVMNED			
g165-1	ELNYAPLGADGVINPARALN	IAEQFHVSQFWATLVAEGK	LEDNSFINAVPHMSLVMNED			
	70	80	90	100	110	120
	130	140	150	160	170	180
m165-1.pep	HCSYLQKRYDAFKTQKLFEN	MEFSTDRNKISDWAPLMMRG	RDENQPVAA	NYSAEQTDVDF		
g165-1	HCRYLQKRYDVFKTQKLFEN	MEFSTDRNKISDWAPLMMRG	RDENQPVAA	NYSAEQTDVDF		
	130	140	150	160	170	180
	190	200	210	220	230	240
m165-1.pep	GRLTRQMVKYLQKGKVKTEF	NRHVEDIKRESDGAWVLKTA	DTRNPDGQLTLTRFLFLGA			
g165-1	GRLTRQMVKYLQKGKVKTEF	NRHVEDIKRESDGAWVLKTA	DTRNPDGQLTLTRFLFLGA			
	190	200	210	220	230	240
	250	260	270	280	290	300
m165-1.pep	GGGALTLLQKSGIPEGKGYG	GPFVSGLFERN	SNPETAQHNAKVYGQASV	GAPPMSVPHL		
g165-1	GGGALTLLQKSGIPEGKGYG	GPFVSGLFERN	SNPETAQHNAKVYGQASV	GAPPMSVPHL		
	250	260	270	280	290	300
	310	320	330	340	350	360
m165-1.pep	DTRNVDGKRHLMFGPYAGFR	SNFLKQGSIMDLPLSIHMDN	LYPMLCAGWANMPLTKYLLG			
g165-1	DTRNVDGKRHLMFGPYAGFR	SNFLKQGSIMDLPLSIHMDN	LYPMLCAGWANMPLTKYLLG			
	310	320	330	340	350	360
	370	380	390	400	410	420
m165-1.pep	ELRKTKEERFASLLEYYPEA	NPDDWELITAGQRVQIHKD	SEKGGVLQFGTEIVAHADGS			
g165-1	ELRKTKEERFASLLEYYPEA	NPDDWELITAGQRVQIHKD	SEKGGVLQFGTEIVAHADGS			
	370	380	390	400	410	420
	430	440	450	460	470	480

475

```

a165-1.pep  GRLTRQMVKYLGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
m165-1      GRLTRQMVKYLGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTLRTRFLFLGA
              190      200      210      220      230      240

              250      260      270      280      290      300
a165-1.pep  GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEOHNAKVYQASVGAPPMSVPHL
m165-1      GGGALTLLQKSGIPEGKGYGGFPVSGLFFRNSNPETAEOHNAKVYQASVGAPPMSVPHL
              250      260      270      280      290      300

              310      320      330      340      350      360
a165-1.pep  DTRNVDGKRHLMFPGPYAGFRSNFLKQGSMDLPLSIHMDNLYPMLRAGWANMPLTKYLLG
m165-1      DTRNVDGKRHLMFPGPYAGFRSNFLKQGSMDLPLSIHMDNLYPMLCAGWANMPLTKYLLG
              310      320      330      340      350      360

              370      380      390      400      410      420
a165-1.pep  ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVAHADGS
m165-1      ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVAHADGS
              370      380      390      400      410      420

              430      440      450      460      470      480
a165-1.pep  LAALLGASPGASTAVPLMIRLMHQCFPERTPSWEGRLKELVPGYGIKLNENPERADEIIA
m165-1      LAALLGASPGASTAVPLMIRLMHQCFPERAPSWEDRLKELVPGYGIKLNENPERADEIIA
              430      440      450      460      470      480

              489
a165-1.pep  YTAKVLDIX
m165-1      YTAKVLDIX

```

**a165-1/p33940**

```

sp|P33940|YOJH_ECOLI HYPOTHETICAL 60.2 KD PROTEIN IN ECO-ALKB INTERGENIC REGION
>gi|1736851|gnl|PID|d1016718 (D90850) ORF ID:o372#5; similar to [SwissProt Accession Number
P33940] [Escherichia coli] >gi|1788539 (AE000310) f548; This 548 aa ORF is 100 pct identical
to 490 residues of YOJH_ECOLI SW: P33940 (492 aa) but contains 56 additional N-ter aa; 100 pct
identical to GB: ECOHU49_33
ACCESSION: U00008 (490 aa) but contains 58 additional N-term resi... Length = 548
Score = 458 bits (1167), Expect = e-128
Identities = 233/490 (47%), Positives = 303/490 (61%), Gaps = 5/490 (1%)

```

```

Query: 3  EATDVVLVGGGIMSATLGVLKLEPSWEITLIERLEDVALESSNAWNNAGTGHSALCEL 62
      + TDV+L+GGGIMSATLG L+ELEP W +T++ERLE VA ESSN WNNAGTGHSAL EL
Sbjct: 30 QETDVLLIGGIMSATLGTYLRELEPEWSMTMVERLEGVAQESSNGWNNAGTGHSALMEL 89

```

```

Query: 63 NYAPLGANGIIDPARALNIAEQFHVSQFQWATLVAEGKLED-NSFINAVPHMSLVMNEDH 121
      NY P A+G I +A+ I E F +SROFWA V G L SFIN VPHMS V ED+
Sbjct: 90 NYTPQNADGSISIEKAVAINAEAFQISQFWAHQVERGVLRTPRSFINTVPHMSFVWGEDN 149

```

```

Query: 122 CSYLQKRYDAFKTQKLFENMEFSTDNRKISDWAPLMMRGRDENQPVAAANYSAEGTDVDFG 181
      ++L+ RY A + LF M +S D +I +WAPL+M GRD Q VAA + GTDV++G
Sbjct: 150 VNFLRARYAALQQSSLFGRMYSEDHAQIKWAPLVMGRDPQQKVAATRTIGTDVNYG 209

```

```

Query: 182 RLTRQMVKYLGKGVKTEFNRHVEDIKRESDGAWVLKTADTRNPDGQLTXXXXXXXXXX 240
      +TRQ++ LQ K + + V +KR D W + AD +N Q
Sbjct: 210 EITRQLIASLQKKSNSFLQLSSEVRALKRNDNTWTVTVADLKNGTAQ-NIRAKFVFIGA 268

```

```

Query: 241 XXXXXXXXQKSGIPEGKGYGGFPVSGLFFRNSNPETAEOHNAKVYQASVGAPPMSVPHL 300
      Q+SGIPE K Y GFPV G F + NP+ H AKVYG+ASVGAPPMSVPH+
Sbjct: 269 GGAALKLLQESGIPEAKDYAGFPVGGQFLVSENPDVVNHHLAKVYGKASVGAPPMSVPHI 328

```

```

Query: 301 DTRNVDGKRHLMFPGPYAGFRSNFLKQGSMDLPLSIHMDNLYPMLRAGWANMPLTKYLLG 360
      DTR +DGKR ++FGP+A F + FLK GSL DL S N+ PM+ G N L KYL+
Sbjct: 329 DTRVLDGKRVLFGPFATFSTKFLKNGSLWDLMSSTTTSNVMPMMHVGLDNFDLVKYLVS 388

```

```

Query: 361 ELRKTKEERFASLLEYYPEANPDDWELITAGQRVQIIKKDSEKGGVLQFGTEIVXXXXXX 420
      ++ ++E+RF +L EYYP+A +DW L AGQRVQIIK+D+EKGGVL+ GTE+V
Sbjct: 389 QVMLSSEDRFEALKEYYPQAKKEDRWLWQAGQRVQIIKRDAEKGGVLRGLGTEVVSQQGT 448

```

```

Query: 421 XXXXXXXXXXXXXVPLMIRLMHQCFPER--TPSWEGRLKELVPGYGIKLNENPERADEI 478

```

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 204 shows 82.0% identity over a 250 aa overlap with a predicted ORF (ORF 204.ng) from *N. gonorrhoeae*:

m204/g204

```

              10      20      30      40      50      60
m204 . pep  MAAAEIKRPFVDFQHIASVLHGGIAAFACLIGLQGGMRNXVIRQFAAVFGDIAHQFGKQ
              |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g204        MAAAEIKRPLAVDFQHIASVLHGGIAAFACLIGLQGGMRNQVISQFAAVFGDIAHQFGKQ
              10      20      30      40      50      60

              70      80      90      100     110     120
m204 . pep  GMAHAVFCPTCRTVLIIGFHTFAADGFQVAGMLADQSDNILFRQAFNRITDLFFAVVGFA
              |||||:|:|:|:|||||:|||||:|||||:|:|:|:|||||:|||||:|||||:|||||
g204        GMAHAVFRPARRRVLSVGFHTFADDGFQVVGMLSGQPDGVLFRQAFNRITDLFFAVVGFA
              70      80      90      100     110     120

              130     140     150     160     170     180
m204 . pep  FAALSQIQTNRRIVDIYDFENRFRRLYRVLRLYRRIXGFAATAXQQAAYGKXXQH
              ||:|||||:|||||:|||||:|||||:|||||:|||||:|:|:|:|||||:|:|
g204        FATLSQSQTGNRRIVDVDFENRFRRLCRILRLFRRIFGFAAGGKQQAAYHGKRYFQH
              130     140     150     160     170     180

              190     200     210     220     230
m204 . pep  STXLMVSKCRLK----RGRRRFGRHRVHFNGRMPTASGTLNNSRASLRAFAAPACKISS
              |:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g204        SALLMVSKCRLKCRLLKRRRRFGRHWVYFNGRMPTASRTLNNNSRASLRAFCAPACKISS
              190     200     210     220     230     240

              240
m204 . pep  ICEGSAVSSLX
              |||||:|:|
g204        ICEGLEVNAL
              250

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 701>:

a204 . seq

```

1  ATGGCGGCGG  CGGAAATAAA  ACGCCCCCTC  GCTGTCGATT  TCCAGCACAT
51  AGCGTCCGTT  CTGCACGGCG  GCATAGCCGC  TTTTGCTGCG  CTGATAGGGT
101  TGCAGGCGCG  AATGCGAAAT  CAGGTAATCC  GTCAGTTTGC  CGCGCTCTTC
151  GGCGATATTG  CCCACCAGTT  TGGCAAACAA  GGTATGGCAC  ACGCCGTTTG
201  CCGCCAGGCC  CGAAGGCGCG  CCCTTCCGT  CGGTTTCCAT  ACATTTGCCG
251  ACGACGGCTT  CCAAGTCGTT  GGGATGCTTG  CCGGTCAGCC  GGACGACGTT
301  TTGTTCCGGC  AAGCCTT...
351  .....
401  .....
451  .....
501  .....
551  .....AAGAG  GTTCGGACGG
601  CATTGGGTTT  ATTTCAACGG  GCGGATACCG  ACCGCATCAC  GTACTTTGCC
651  CAATAATTGG  CGTGCTTCTT  TACGCGCTTT  TTGCGCGCCT  GCCTGCAAAA
701  TCTCTTCGAT  TTGCGAAGGG  TCGCGGTGTA  GCTCGTTGTA  G

```

This corresponds to the amino acid sequence <SEQ ID 702; ORF 204.a>:

a204 . pep

```

1  MAAAEIKRPL  AVDFQHIASV  LHGGIAAFAC  LIGLQGGMRN  QVIRQFAAVF
51  GDIAHQFGKQ  GMAHAVCRPA  RRRALSVGFH  TFADDGFQVV  GMLAGQPDVV
101  LFRQAF....
151  .....KRFRGR
201  HWVYFNGRIP  TASRTLNNNS  RASLRAFCAP  ACKISSICEG  SAVSSL*

```



This corresponds to the amino acid sequence <SEQ ID 706; ORF 205>:

```
m205.pep
1  MLXTXFAVLG GCLLXCRCGK SXNTAVQPQN AVQSAPKPVF KVIYIDNTAI
51  AGLDLGQSSE GKTNDGKKQI SYPIKGLPEQ NVIRLIGKHP GDLEAVSGKC
101 METDDKDSPA GWAENGVCHT LFAKLVGNIA EDGGKLTDL VSHAALQPYQ
151 AGKSGYAAVQ NGRYVLEIDS EGAFYFRRRH Y*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 205 shows 88.4% identity over a 181 aa overlap with a predicted ORF (ORF 205.ng) from *N. gonorrhoeae*:

```
m205/g205

      10      20      30      40      50      60
m205.pep  MLXTXFAVLGGCLLXCRCGKSXNTAVQPQNAVQSAPKPVFKVIYIDNTAIAGLDLGQSSE
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
g205      MLKIPFAVLGGCLLLAACGKSENTAEQPQNAASAPKPVFKVKYIDNTAIAGLALGQSSE
          10      20      30      40      50      60

      70      80      90      100     110     120
m205.pep  GKTNDGKKQISYPIKGLPEQNVIRLIGKHPGDLEAVSGKCMETDDKDSPA GWAENGVCHT
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
g205      GKTNDGKKQISYPIKGLPEQNAVRLTGKHPNDLEAVVGKCMETDGKDAPSGWAENGVCHT
          70      80      90      100     110     120

      130     140     150     160     170     180
m205.pep  LFAKLVGNIAEDGGKLTDLVSHAALQPYQAGKSGYAAVQNGRYVLEIDSEGAFFYFRRRH
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
g205      LFAKLVGNIAEDGGKLTDLVSHSALQPYQAGKSGYAAVQNGRYVLEIDSEGAFFYFRRRH
          130     140     150     160     170     180

m205.pep  YX
          |
g205      Y
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 707>:

```
a205.seq (partial)
1  TCCGAACCTC TTAAAGGCTT GCCGGAACAA AACGTCGTCC GGCTGACCGG
51  CAAGCATCCC AACGACTTGG AAGCCGTCGT CGGCAAATGT ATGGAACACG
101 ACGGAAAGGG CGCGCCTTCG GGCTGGGCGG CAAACGGCGT GTGCCATACC
151 TTGTTTGCCA AACTGGTGGG CAATATCGCC GAAGACGGCG GCAAACTGAC
201 GGATTACCTG ATTTTCGATT CCGCCCTGCA ACCCTATCAG GCAGGCAAAA
251 GCGGCTATGC CGCCGTGCAG AACGGACGCT ATGTGCTGGA AATCGACAGC
301 GAGGGGGCGT TTTATTTCG CCGCCGCCAT TATTGA
```

This corresponds to the amino acid sequence <SEQ ID 708; ORF 205.a>:

```
a205.pep (partial)
1  SEPLKGLPEQ NVVRLTGKHP NDLEAVVGKC METDGKGAPS GWAANGVCHT
51  LFAKLVGNIA EDGGKLTDL ISHSALQPYQ AGKSGYAAVQ NGRYVLEIDS
101 EGAFYFRRRH Y*
```

**m205/a205** 88.3% identity in 111 aa overlap

```
      50      60      70      80      90      100
m205.pep  KVIYIDNTAIAGLDLGQSSEGKTNDGKKQISYPIKGLPEQNVIRLIGKHPGDLEAVSGKC
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a205      SEPLKGLPEQNVVRLTGKHPNDLEAVVGKC
          10      20      30

      110     120     130     140     150     160
m205.pep  METDDKDSPA GWAENGVCHTLFAKLVGNIAEDGGKLTDLVSHAALQPYQAGKSGYAAVQ
```

481

```

              70      80      90      100      110      120
          130      140      150      160      170
g205-1.pep  LFAKLVGNI AEDGGKLT DYLI SHSALQPYQAGKSGYAAVQNGRYVLEIDSEGAF
m205-1      LFAKLVGNI AEDGGKLT DYLV SHAALQPYQAGKSGYAAVQNGRYVLEIDSEGAFYFRRRH
          130      140      150      160      170      180
m205-1      YX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 713>:

```

a205-1.seq (partial)
1   CCTCTTAAAG GCTTGCCGGA ACAAACGTC GTCCGGCTGA CCGGCAAGCA
51  TCCCAACGAC TTGGAAGCCG TCGTCGGCAA ATGTATGGAA ACCGACGGAA
101 AGGGCGCGCC TTCGGGCTGG GCGGCAAACG GCGTGTGCCA TACCTTGTTC
151 GCCAACTGG TGGGCAATAT CGCCGAAGAC GCGGCAAAAC TGACGGATTA
201 CCTGATTTCG CATTCCGCC TGAACCCCTA TCAGGCAGGC AAAAGCGGCT
251 ATGCCGCCGT GCAGAACGGA CGCTATGTGC TGGAAATCGA CAGCGAGGGG
301 GCGTTTATT TCCGCCGCC CCATTATTGA

```

This corresponds to the amino acid sequence <SEQ ID 714; ORF 205-1.a>:

```

a205-1.pep (partial)
1   PLKGLPEQNV VRLTGKHPND LEAVVGKCM ETDGKGAPSGW AANGVCHTLF
51  AKLVGNI AEDGGKLT DYLI SHSALQPYQAG KSGYAAVQNG RYVLEIDSEG
101 AFYFRRRH*

```

m205-1/a205-1 89.0% identity in 109 aa overlap

```

          50      60      70      80      90      100
m205-1.pep KYIDNTAIAGLDLGQSSEGKTNDGKKQISYPIKGLPEQNVIRLIGKHPGDLEAVSGKCM E
          |||||
a205-1      PLKGLPEQNVVRLTGKHPNDLEAVVGKCM E
          10      20      30

          110      120      130      140      150      160
m205-1.pep TDDKDSPAGWAENGVCHTLFAKLVGNI AEDGGKLT DYLV SHAALQPYQAGKSGYAAVQNG
          || :||| |||||
a205-1      TDGKGAPSGWAANGVCHTLFAKLVGNI AEDGGKLT DYLI SHSALQPYQAGKSGYAAVQNG
          40      50      60      70      80      90

          170      180
m205-1.pep RYVLEIDSEGAFYFRRRH YX
          |||||
a205-1      RYVLEIDSEGAFYFRRRH YX
          100      110

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 715>:

```

g206.seq
1   atgttttccc ccgacaaaac ccttttcttc tgtctcgggc cactgtctct
51  cgctcatgc ggcacgacct ccggcaaaaca ccgccaaccg aaacccaaac
101 agacagtccg gcaaatccaa gccgtccgca tcagccacat cggccgcaca
151 caaggctcgc aggaactcat gctccacagc ctggactca tcggcacgcc
201 ctacaaatgg ggcggcagca gcaccgcaac cggcttcgac tgcagcgga
251 tgattcaatt ggtttacaaa aacgccctca acgtcaagct gccgcgcacc
301 gcccgcgaca tggcggcggc aagccgcaaa atccccgaca gccgcctcaa
351 ggccggcgac atcgattctc tcaacaccgg cggcgcacac cgctactcac
401 acgtcggact ctacatcggc aacggcggaat tcatccatgc ccccgcgagc
451 ggcaaaacca tcaaaaccga aaaactctcc acaccgtttt acgccaaaaa
501 ctaccttga gcgcatacgt tttttacaga atga

```

This corresponds to the amino acid sequence <SEQ ID 716; ORF 206.ng>:

```

g206.pep
1   MFSPDKTLFL CLGALLLASC GTTSGKHRP KPKQTVRQIQ AVRISHIGRT

```

483

This corresponds to the amino acid sequence <SEQ ID 720; ORF 206.a>:

```
a206.pep
  1 MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
 51 QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPR
101 ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 99.4% identity in 177 aa overlap

	10	20	30	40	50	60
m206.pep	MFPPDKTLFLCLSALLLASC	GTTSGKHRQPKPKQTVRQIQ	AVRISHIDRTQGSQELMLHS			
a206	MFPPDKTLFLCLSALLLASC	GTTSGKHRQPKPKQTVRQIQ	AVRISHIDRTQGSQELMLHS			
	10	20	30	40	50	60
	70	80	90	100	110	120
m206.pep	LGLIGTPYKWGGSSTATGFD	CSGMIQFVYKNALNVKLPR	TARDMAAASRKIPDSRXKAGD			
a206	LGLIGTPYKWGGSSTATGFD	CSGMIQFVYKNALNVKLPR	TARDMAAASRKIPDSRLKAGD			
	70	80	90	100	110	120
	130	140	150	160	170	
m206.pep	LVFFNTGGAHRYSHVGLYIG	NGEFIHAPSSGKTIKTEKLS	TPFYAKNYLGAHTFFTEX			
a206	LVFFNTGGAHRYSHVGLYIG	NGEFIHAPSSGKTIKTEKLS	TPFYAKNYLGAHTFFTEX			
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 721>:

```
g209.seq
  1 atgctgcggc atttaggaaa cgccttgcgc ttgggcgcgt tgtttttcga
 51 tgctgcgggt gatgtgccac tgctgggcga tggtcaggag gttgttgacc
101 acccagtaga gaaccaaacc ggcagggaag aagaagaaca tgacggagaa
151 aaccaacggc atgattttca tcattttcgc ctgcatcggg tcggtcggcg
201 gcgggttcag ataggtttg gcgaacatcg ttgccgccat aatgatgggc
251 aggatgtagt aggggtcggc gcggctgagg tcggtaatcc agcccagcca
301 aggtgcctgg cgcaattcta cggaggcgaa caatgccacg tacaagccga
351 tgaagacggg gatttgcaac agcataggca gacagccgcc cagcgggttg
401 atttcctcgt ctgcgaaaag ctgcatcacc gcttgctgtt gcgccatacg
451 gtcgtcgccg tatttttctt tgatggtctg cagttcgggt gcggcgccac
501 gcattttcgc catcgaaagg taggaggcgt tggtaaatgg atacagtacg
551 gctttgacga tgatggtcaa aacgacgatt gcccagcccc agttgccgat
601 aatgttgatg agttggttca ggagccagaa gagcgcgcat gcgaaccagt
651 gtactttacc gtagtctttt gccagttgca ggttgctcggc gatgtttgcg
701 ataacggatg tggtttcggg accggcatac aggttgaccg ccattttcgg
751 ttttgcccc cgggttgga tagcgggttaa
```

This corresponds to the amino acid sequence <SEQ ID 722; ORF 209.ng>:

```
g209.pep
  1 MLRHLGNDEA LGALFFDAAV DVPLLGDGQE VVDHPVENQT GREEEHDEGE
 51 NQRHDFHHFR LHRVGRRRVQ IGLGEHRCRH NDQDQVVGVG AAEVGNPAQP
101 RCLAQFYGGE QCPVQADEDG DLQQRQTAA QRVDLFVEK LHHRLLLRHT
151 VVAVFFFDGL QFGCGGTHFR HRTVGGVGQW IQYGFDDDGQ NDDCPAPVAD
201 NVVQLVQEP EERRCEPVYFT VFECQLQVVG DVCDNGCGLR TGIQVDRHER
251 FWPPGWDSG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 723>:

```
m209.seq
  1 ATGCTGCGGC ATTTAGGAAA CGACTTCGCC TTGGGCGGTT GTTTTTCGAT
 51 GCTGCGGTTG ATGTGCCATT GCTGGGCGAT GGTCAGGAGG TTGTTGACTA
101 CCCAGTACAA TACCAGACCG GCAGGGAAGA AGAAGAACAT GACGGAGAAA
151 ACCAACGGCA TGATTTTCAT CATTTCGCC TGCATCGGT CGGTGCGCGG
201 CGGGTTCAGA TAAGTTTGGG CGAACATCGT TGCCGCCATA ATGATGGGCA
251 GGATGTAGTA GGGGTCGGCG CGGCTGAGGT CGGTAATCCA ACCCAGCCAA
```

485

```

51  TGCTGCGGTT GATGTGCCAT TGCTGGGCGA TGGTCACGAG GTTGTTGATC
101  ACCCAGTACA ATACCAGACC GGCAGGGAAG AAGAAGAACA TGACGGAGAA
151  AACCAGAGGC ATGATTTTCA TCATTTTCGC CTGCATCGGG TCGGTCGGCG
201  GCGGGTTTCA ATAGGTTTGG GCGAACATCG TTGCCGCCAT AATGATGGGC
251  AGGATGTAGT AGGGGTTCGC GCGGCTGAGG TCGGTAATCC AATCCAGCCA
301  AGGTGCCTGG CGCAATTCTA CGGAGGCGAA CAATGCCCAA TACAATCCGA
351  TGAAGACGGG GATTTGCAAC AGCATAGGCA GGCAGCCGCC CAGCGGGTTG
401  ATTTTCTCGT CTGTGTAAAG CTGCATCATG GCTTGTGTCT GCGCCATACG
451  GTCGTGCGCG TATTTCTCTT TGATGGCTTG CAGTTTGGGC GCGGCGGCAC
501  GCATTTTCGC CATCGAACGG TAAGAGGCGT TGGTCAATGG ATACAGTACG
551  GCTTTGACGA TGATGGTTAA AACGATAATC GCCCAGCCCC AGTTGCCGAT
601  GATGTTGTGC AGTTGGTTCA AAAGCCAAAA GAGGCGGAG GCGAACCACT
651  GTACTTTGCC GTAGTCTTTG GCCAGTTGCA GGTGTGCGGC GATGTTTGGC
701  ATAACGGATG TGGTCTGTGG GCCGGCGTAG AGGTTGATGG AGGCTTCGGT
751  TTCGCACCGT TTTGGATAGC GGCTAAAGGC ACGCTGACGC TGGTGTCTGA
801  CAGCTTGTGC TTGCGGCGTT TGATGTCGAT ACGGCAGTCG CCAGCGGCGC
851  AAACGCTTTG TCCGCTTTT GGTTCGAGGA TCCAGGTGGA CATGAAGTGG
901  TGTTCAATCA TGCCGAGCCA GCCGGTCGGG GTTTTCCGGA TGTATTGCGC
951  CTCGGATTGG CCGGATTGG CATCGTCGTC CAAGTCGGAG AAGCTGACTT
1001 TTTGGAAGTT GCCTTCAGGG GTATAA

```

This corresponds to the amino acid sequence <SEQ ID 726; ORF 209.a>:

```

a209.pep
1  MLRHLGNDEA LGALFFDAAV DVPLLGDGQE VVDHPVQYQT GREEEHDGE
51  NQRHDFHHFR LHRVGRRRVQ IGLGEHRCRH NDGQDVVGVG AAEVGNPTQP
101 RCLAQFYGGE QCPIQSDGDL DLQHRQAAA QRVDFLVCVK LHHGLLRHT
151 VVAVFLFDGL QFGRGGTHFR HRTVRGVQW IQYGFDDG* NDNRPAPVAD
201 DVVQLVQPK EGGGEPVYFA VVFGQLQVVG DCDNGCGLW AGVEVDGGFG
251 FAPFWIAAKG TLTVLVLSL LRLMSIRQS PAAQTLCPPL GWRIQVDMKW
301 CSIMPSQPVG VLRMYSASDL PDLASSSKSE KLTFWKLPSG V*

```

m209/a209 95.6% identity in 341 aa overlap

	10	20	30	40	50	60
m209.pep	MLRHLGNDFALGALFFDAAVDVPLLGDGQEVVDYVPVQYQTG	EEEEHDGENQRHDFHHFR				
a209	MLRHLGNDFALGALFFDAAVDVPLLGDGQEVVDHPVQYQTG	EEEEHDGENQRHDFHHFR				
	10	20	30	40	50	60
	70	80	90	100	110	120
m209.pep	LHRVGRRRVQISLGEHRCRHNDGQDVVGVA	AAEVGNPTQPRCLAQFYGGEQCPIQSDG				
a209	LHRVGRRRVQISLGEHRCRHNDGQDVVGVA	AAEVGNPTQPRCLAQFYGGEQCPIQSDG				
	70	80	90	100	110	120
	130	140	150	160	170	180
m209.pep	DLQHRQAAAQRVDFLVCVKLHHRLLLRHTVVAVFLFDGLQF	CGGTHFRHRAVRGVGQW				
a209	DLQHRQAAAQRVDFLVCVKLHHRLLLRHTVVAVFLFDGLQF	GRGTHFRHRTVRGVGQW				
	130	140	150	160	170	180
	190	200	210	220	230	240
m209.pep	IQYGFDDGDXNDNRPA	PVADDVVQLVQEP	EEGGEPVYFAVVFGQLQVVG	DCDDGCGGLR		
a209	IQYGFDDGDXNDNRPA	PVADDVVQLVQEP	EEGGEPVYFAVVFGQLQVVG	DCDNGCGLW		
	190	200	210	220	230	240
	250	260	270	280	290	300
m209.pep	AGVEVDGGFGFAPFWMAAKGTLTLVLVLSL	LRLMSMLHSPAAQTLCLPLGWRIQVDMKW				
a209	AGVEVDGGFGFAPFWIAAKGTLTLVLVLSL	LRLMSIRQSPAAQTLCPPLGWRIQVDMKW				
	250	260	270	280	290	300
	310	320	330	340		
m209.pep	CSIMPSQPVGVLRMYSASDL	PDLASSSKSEKLTFWKLPSGVX				

g211	AFVVLQREVTFFGEDDVVAAVFLVFAQEQVPHDAGFGVKGFDEINPAVALAQVIELAGFH	70	80	90	100	110	120
m211.pep	QRQFLLLQLQDFSVFAAAXLCPRYHPKLHDGNGQNGKRHGKLLHHRAYPLFQCQSAGX	130	140	150	160	170	
g211	QRQFLLLQLQNFGIFAAALCPRYHPKLHDGNGQDKRHGKLLHDGAYPLFQORQSAG	130	140	150	160	170	

```

a211.seq
1  ATGTTGCGGG TTGCTGCTGC CAATCAGTTG GCGGTCGGA ATGGTACGGC
51  GGTGGGAAAC GGGGTCGATG AGTTTGGGCG TGGTGCTGAT AATCAGGTTG
101 AGTTTPTTGA AGAAACCTCG ATTTGATGTC GCGCGCTCCG GCGGTGCCCT
151 GTAACGGTAG CCGTGCGCA ATTCGAGCGT GCGTTTGTG TCGTTCAGCG
201 AGAAGTTACT TTCTTTGGCG AAGATGATGT TGTGCGCCGC GTTTTGTGCC
251 TGTTTCGCGA TGAACAGGTT TTTCATGATG CCGGATTCGG TATCGAAGTG
301 TTCGACAAAA GAAACCTGCG CGTTGCGGCT CCGCAAACTG TTGAACCGCG
351 CTGCCCTCCAC CAAAGACAAT TCCTGCTTCT GCTTCAGGAT TTCAGCGTAT
401 TCGCGGCTGC GTAGCTCTGC CCAGCGATAT ACCCAAAGCT GCATGACGGC
451 AACCAAAACG CAACACGGCA CCGGAAACTG CATCACGGG CGTATCCATT
501 GTTTCAAATG CAAATCCGAG GATAG

```

a211.pep

1	MLRVAANQL	GGRNGTAVGN	GVDEFGRGAD	NQVEFLEGNL	<u>IVVGASGRAA</u>
51	VTVAVAQFER	AFVVVQREVT	FFGEDDVVAA	VFVLFAQEQV	FHDAGFGIEG
101	FDKINPAVAL	AQTVEPACLR	<u>QRQFLLLQD</u>	<u>FSVFAAA</u>	<u>LC</u> PRYHPKLHDG
151	NONGKRHGKL	HHRAYPLFOC	OSAG*		

	10	20	30	40	50	60
m211.pep	MLRVAAANQLGGRNGTAVGNGVDEFGRGADNQVEFLEGNLI	VVGASGRAAVTVAVAQFER				
a211	MLRVAAANQLGGRNGTAVGNGVDEFGRGADNQVEFLEGNLI	VVGASGRAAVTVAVAQFER				
	10	20	30	40	50	60
	70	80	90	100	110	120
m211.pep	AFVVVQREVTFFGEDDVVAAVFVLF	FAQEQVFHDAGFGIEGFDKINPAVALAQTVELAC				
a211	AFVVVQREVTFFGEDDVVAAVFVLF	FAQEQVFHDAGFGIEGFDKINPAVALAQTVEPAC				
	70	80	90	100	110	120
	130	140	150	160	170	
m211.pep	QRQFLLLLQDFS	VFAAAXLCPRYHPKLHDG	NQNGKRHGK	LHHRAYPL	FQCQSAGX	
a211	QRQFLLLLQDFS	VFAAAXLCPRYHPKLHDG	NQNGKRHGK	LHHRAYPL	FQCQSAGX	
	130	140	150	160	170	

```
g212.seq (partial)
      1  atggacaatc  tctgatggga  cggcattccc  gacatccgca  cactcgacca
     51  aaccatccgc  aaacacgcac  acccgctcaa  cctgattgtc  tgectccccg
    101  ataatacagat  tcccgatttt  caaaccgcaa  aagatgcttc  ggactcgga
    151  tgcgctetga  agcaccgttt  ggatcaggca  acccagtgcc  tccagttcga
    201  cagcatcaac  ctcatcgaac  acatcctgcc  cgatgtccgc  ttctggctgg
    251  tcccccttc  aacgacccgc  cgctgcag  aacacttcca  ccacatttcc
    301  ttggcacacg  agccattccc  gcaaacggaa  agcaagtcga  acaaacctg
    351  gtttgactt  ccacaaacat  ccgaacggaa  aaaaacggaa  cagctcctcg
```

1501 GmCAwTGCAG CCCAAATCst AGGCyTGCCC CATCCCTTTT yAcAAcGCCT  
 1551 gCGCCACGCC cTAcACCCCA ACCGCACCAT CATCCGCGCC ATCGTCAGAA  
 1601 GGAAGGATCT AACCCTTAA

This corresponds to the amino acid sequence <SEQ ID 736; ORF 212>:

m212.pep  
 1 MDNLVWDGIP DIRTLDQAIR KHAPPLNLII CLPDNQIPDF QTAQDASDAE  
 51 CRLKHRLDQA MQCLQFDSIN LIEHILPDVR FWLVPPSRTH HLHEHFHHIS  
 101 WQTEAIPQTE SKPDKPWFAL POTSERQKPE HILVIGAGIS GAATAHALAS  
 151 HGISVTVLEA RKAAQAASGN RQGLLYAKIS PHDTEQTELL LAGYGYTKRL  
 201 LGHILPESET WGGNGIIHLN YSRTEQQRNH ELGLQKHNNH LYRSITSAEA  
 251 EKIAGIPLSV PYDHPSCGLY WQHGVLNPP AFVRTLLNHP LIGLHEDTPL  
 301 TDISHDGEKW IASTPNGTFT ATHIYCTGA NSPYLPETNL AALPLRQIRG  
 351 QTGLTPSTPF SEQLRCAVSG ESYISPSWHG LHCYGASFI PNSSHTGWN  
 401 EEASNRQALA HLNPALESSE FAANPNPKH QGHAAIRCD PDHLPLVGAL  
 451 GDIAAMRQTY TKLALDKNYR IDTPCPYLPN AYVNTAHGTR GLATAPICAA  
 501 XXAAQIXGLP HFFXQRLRHA LHPNRTIIRA IVRRKDLTP\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 212 shows 92.9% identity over a 421 aa overlap with a predicted ORF (ORF 212.ng) from *N. gonorrhoeae*:

m212/g212

	10	20	30	40	50	60
m212.pep	MDNLVWDGIPDIRTLDQAIRKHAPPLNLII	CLPDNQIPDFQTAQDASDAECRLKHRLDQA				
g212	MDNLVWDGIPDIRTLDQTIRKHAHPLNLIV	CLPDNQIPDFQTAQDASDSECRKHLRDLQA				
	10	20	30	40	50	60
	70	80	90	100	110	120
m212.pep	MQCLQFDSINLIEHILPDVRFWLVPPSRTH	HLHEHFHHISWQTEAIPQTESKPKPWFAL				
g212	TQCLQFDSINLIEHILPDVRFWLVPPSRTR	RLHEHFHHISWQTEAIPQTESKSDKPWFAL				
	70	80	90	100	110	120
	130	140	150	160	170	180
m212.pep	POTSERQKPEHILVIGAGISGAATAHALASH	GISVTVLEARKAAQAASGNRQGLLYAKIS				
g212	POTSERKKPEHVLVIGAGIAGASTAHALASH	GISVTVLEARKAAQAASGNRQGLLYAKIS				
	130	140	150	160	170	180
	190	200	210	220	230	240
m212.pep	PHDTEQTELLLAGYGYTKRLLGHILPESET	WGGNGIIHLNYSRTEQQRNH	ELGLQKHNNH			
g212	PHDTGQTELLLAGYGYTKRLLGHILPDS	TWGGNGIIHLNYSRTEQQRNH	ELGLQKHNNH			
	190	200	210	220	230	240
	250	260	270	280	290	300
m212.pep	LYRSITSAEAEKIAGIPLSVPYDHPSCGLY	WQHGVLNPPAFVRTLLNHP	LIGLHEDTPL			
g212	LYRSITSAEAEKIAGIPLNTPYAEPLCGLY	WQHGVLNPPAFVRTLLSHPLIELYENTTL				
	250	260	270	280	290	300
	310	320	330	340	350	360
m212.pep	TDISHDGEKWI	ASTPNGTFTATHI	IYCTGANS	PYLPETNL	AALPLRQIRG	QTGLTPSTPF
g212	TGISHDGEKWI	ASTPNGTFTATHI	IYCTGAHSP	CLPETNL	AALPLRQIRG	QTGLTPSTPF
	310	320	330	340	350	360
	370	380	390	400	410	420
m212.pep	SEQLRCAVSGESYISPSWHGLHCYGASFI	PNSSHTGWN	EEASNRQALAH	LNPALESSE		

491

	10	20	30	40	50	60
	70	80	90	100	110	120
m212.pep	MQCLQFDSINLIEHILPDVRFWLVPSPRTHLHEHEFHHSWQTEAIPQTESKPKDPWFAL					
a212	TQCLQFDSINLIEHILPDVRFWLVPSPRTHLHEHEFHHSWQTEAIPQTESKPKDPWFAL					
	130	140	150	160	170	180
m212.pep	PQTSERQKPEHILVIGAGISGAATAHALASHGISVTVLEARKAAQAASGNRQGLLYAKIS					
a212	PQTSERQKPEHILVIGAGISGAATAHALASYGISVTVLEARKAAQAASGNRQGLLYAKIS					
	190	200	210	220	230	240
m212.pep	PHDTEQTELLLAGYGYTKRLLGHILPESETWGGNGIIHLNYSRTEQQRNHELGLQKHHNH					
a212	PHDTEQTELLLAGYGYTKRLLGHILPESETWGGNGIIHLNYSRTEQQRNHELGLQKHHNH					
	250	260	270	280	290	300
m212.pep	LYRSITSAAEAKIAGIPLSVPYDHPSCGLYWQHGVLNPPAFVRTLNLNHLIGLHEDTPL					
a212	LYRSITQAAEAKIAGIPLNTPYAEPLCGLFWQYGVWLNPPFVRRLLSHPLIGLHEDTPL					
	310	320	330	340	350	360
m212.pep	TDISHDGEKWIASTPNGTFTATHIYCTGANSPLYLPETNLALPLRQIRGQTGLTFSTPF					
a212	TDISHDGEKWIASTPNGTFTATHIYCTGANSPLYLPETNLATLPLRQIRGQTGLTFSTPF					
	370	380	390	400	410	420
m212.pep	SEQLRCAVSGESYISPSWHGLHCYGASFIPNSSHTGWNEAEASNRQALAHLPALSES					
a212	SEQLRCAVSGESYISPSWHGLHCYGASFIPNSSHTGWNEAEASNRQALAHLPALSES					
	430	440	450	460	470	480
m212.pep	FAANPNPQKHQGHAAIRCDSPDHLPLVGALGDIAAMRQTYTKLALDKNYRIDTPCPYLPN					
a212	FAANPNPQKHQGHAAIRCDSPDHLPLVGALGDIAAMQTYAKLALDKNYRIDAPCPYLPN					
	490	500	510	520	530	540
m212.pep	AYVNTAHGTRGLATAPICAAXXAAQIXGLPHFPXQRLRHALHPNRTIIRAIVRRKDLTPX					
a212	AYANTAHGTRGLATAPICAAVAAEILGLPHPLSKRLRHALHPNRAIIRAIVRRKDLTPX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 739>:

```

g214.seq
1  atgatacaaa agatatgtaa gctatttggt ttaattgtaa tttttgcaac
51  ttctcccgtt tttgcccttc aaagcgacag cagacggccc atccaaatcg
101 aagccgacca aggttcgctc gatcaagcca accaaaggac cacatttagc
151 ggcaatgtca tcatcagaca gggtagcgtc aacatttcgg cctcgtgtgt
201 caacgtcaca cgcggcaggc aaaggcggcg aatccgtgag ggcggaaggt
251 tcgcccgtcc gcttcagcca aacggtggac gggggcaaag ggacggtgcg
301 cggtcaggca aacaacgtta cctattcctc cgcagggaagc actgtcgttc
351 tgaccggcaa tgccaaagtg cagcgcggtc gcgacgttgc cgaagggtgc
401 gtcattacct acaacaccaa aaccgaagtc tataccatca acgcgagcac
451 gaaatcgggt gcgaaatccg ctccaaaac cggcagggtc agcgtcgtca
501 tccagccttc aagcacacaa aaaaccgaat aaccccgatg ccgtctgaaa
551 cggaaacgca gttcagacgg catttgccga ccgaaatgcc gagaagagat
601 tattga

```

This corresponds to the amino acid sequence <SEQ ID 740; ORF 214.ng>:

```

g214.pep

```

493

```

351 TAACCGGTAA TGCCAAAGTA CAGCGCGGCG GCGATGTGCG CGAAGGTGCG
401 GTGATTACAT ACAACACCAA AACCGAAGTC TATACCATCA GCGGCAGCAC
451 AAAATCCGGC GCAAATCCG CTTCCAAATC CGGCAGGGTC AGCGTCGTTA
501 TCCAGCCTTC GAGTACGCAA AAATCCGAAT AATCCCAATG CCGTCTGAAA
551 CATAAACCTG GTTCGGACGG CATTGCGCGA CCGAAATATT GAAGAGATAT
601 TTATGA

```

This corresponds to the amino acid sequence <SEQ ID 744; ORF 214.a>:

```

a214.pep
1  MIQKICKLFV LIAFFSASPA FALQSDSRQP IQIEADQGS L DQANQSTTFS
51  GNVVIRQGT L NISAAARNVT RGXQRRRIE GGRFASPLQ P DIGRRQRHGA
101 RTGKQRCFL I CROHRS LNR* CQSTARRRC RRCGDIYQH Q NRSLYHQRH
151 KIRRKIRFQ I RQQRKYP AF EYAKIRIIP M PSET*TWFG R HLPTEILKRY
201 L*

```

m214/a214 99.3% identity in 152 aa overlap

	10	20	30	40	50	60
m214.pep	MIQKICKLFV	LIAFFSASPA	FALQSDSRQP	IQIEADQGS L	DQANQSTTFS	GNVVIRQGT L
a214	MIQKICKLFV	LIAFFSASPA	FALQSDSRQP	IQIEADQGS L	DQANQSTTFS	GNVVIRQGT L
	10	20	30	40	50	60
	70	80	90	100	110	120
m214.pep	NISAAARNVT	RGXQRRRIE	GGRFASPLQ	PDIGRRQRH	GARTGKQRC	FLICRQHRS LNRX
a214	NISAAARNVT	RGXQRRRIE	GGRFASPLQ	PDIGRRQRH	GARTGKQRC	FLICRQHRS LNRX
	70	80	90	100	110	120
	130	140	150			
m214.pep	CQSTARRRC	RRRCGDIYQH	QNRSLYHQ	RHQH KI		
a214	CQSTARRRC	RRRCGDIYQH	QNRSLYHQ	RHQH KIRRKIRFQIR	QGRRYP AF EYAKIRIIP M	
	130	140	150	160	170	180
a214	PSETXTWFG	RHLPT EILKRYLX				
	190	200				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 745>:

```

g214-1.seq
1  ATGATACAAA AGATATGTAA GCTATTTGTT TTAATTGTAA TTTTTCGAAC
51  TTCTCCCGCT TTTGCCCTTC AAAGCGACAG CAGACGGCCC ATCCAAATCG
101 AAGCCGACCA AGGTCGCTC GATCAAGCCA ACCAAAGTAC CACATTTAGC
151 GGCAATGTCA TCATCAGACA GGTACGCTC AACATTTCCG CCTCGCGCGT
201 CAACGTCACA CGCGCGGCA AAGCGGCGA ATCCGTGAGG GCGGAAGGTT
251 CGCCGTCCG CTTACGCCAA ACCTTGGACG GGGGCAAAGG GACGGTCCGC
301 GGTACGGCAA ACAACGTTAC CTATTCCTCC GCAGGAAGCA CCGTCGTTCT
351 GACCGGCAAT GCCAAAGTGC AGCGCGGCGG CGACGTTGCC GAAGGTGCGG
401 TCATTACCTA CAACACCAA ACCGAAGTCT ATACCATCAA CGGCAGCACG
451 AAATCGGTG CGAAATCCGC TTCAAACC GGCAGGGTCA GCGTCGTCAT
501 CCAGCCTTCA AGCACACAAA AAACCGAATA A

```

This corresponds to the amino acid sequence <SEQ ID 746; ORF 214-1.ng>:

```

g214-1.pep
1  MIQKICKLFV LIVIFATSPA FALQSDSRRP IQIEADQGS L DQANQSTTFS
51  GNVVIRQGT L NISASRVNT RGGKGGSVR AEGSPVRF SQ TLDGKGTVR
101 GQANNVTYSS AGSTVLTGN AKVQRGGDVA EGAVITYNT K TEVYTINGST
151 KSGAKSASKT GRVSVVIQPS STQKTE*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 747>:

```

m214-1.seq
1  ATGATACAAA AGATATGTAA GCTATTTGTT TTAATAGCAT TTTTTCGGC
51  GTCCCCCGCT TTTGCCCTTC AAAGCGACAG CAGGCAGCCT ATTCAGATTG
101 AGGCCGACCA AGGTCGCTC GATCAAGCCA ACCAAAGCAC CACATTAGC
151 GGAACGCTG TCATCAGACA GGTACGCTC AATATTTCCG CCGCCCGCGT

```



495

```

51  GNVVIRQGTI NISAARVNVT RGGKGGESVR AEGSPVRFSSQ TLDGGKGTVR
101  GOANNVAYSS AGSTVVLTGN AKVQRGGDVA EGAVITYNTK TEVYTISGST
151  KSGAKSASKS GRVSVVIQPS STQKSE*

```

a214-1/m214-1 100.0% identity in 176 aa overlap

```

          10      20      30      40      50      60
a214-1.pep MIQRICKLFVLIAFFSASPALQSDSRQPIQIEADQSLDQANQSTTFSGNVVIRQGTI
          |||
m214-1      MIQRICKLFVLIAFFSASPALQSDSRQPIQIEADQSLDQANQSTTFSGNVVIRQGTI
          10      20      30      40      50      60

          70      80      90     100     110     120
a214-1.pep NISAARVNVTTRGGKGGESVRAEGSPVRFSSQTLDDGGKGTVRGQANNVAYSSAGSTVVLTGN
          |||
m214-1      NISAARVNVTTRGGKGGESVRAEGSPVRFSSQTLDDGGKGTVRGQANNVAYSSAGSTVVLTGN
          70      80      90     100     110     120

          130     140     150     160     170
a214-1.pep AKVQRGGDVAEGAVITYNTKTEVYTISGSTKSGAKSASKSGRVSVVIQPSSTQKSEX
          |||
m214-1      AKVQRGGDVAEGAVITYNTKTEVYTISGSTKSGAKSASKSGRVSVVIQPSSTQKSEX
          130     140     150     160     170

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 751>:

```

g215.seq
1  atgaaagtaa gatggcggtta cggaattgcg ttccattga tattggcggg
51  tgccttgggc agcctgtcgg catggttggg ccgtaacagc gaagtcgaaa
101  tcgaggaagt caggtcfaat cccgacgaac ctcaatacac aatggacggc
151  ttggacggaa ggcggtttga cgaacaggga tacttgaaag aacatttgag
201  cgcgaaaggt gcgaaacagt ttcccgaaaa cagcgacatc cattttgatt
251  cgccgcatct cgtgttcttc caagaaggca ggctgttgta cgaagtcggc
301  agcgatgaag ccgtttacca taccgaaaac aaacagggtt tttttaaaaa
351  caacgttggtg ctgacaaaaa ccgccgacgg caggcggcag gcgggtaaaag
401  tcgaaaccga aaaactgcac gtcgataccg aatctcaata tgcccaaac
451  gatacgctcg tcagtttcca atatggcgcg tcgcacgggc aggcggcgcg
501  tatgacctac aaccacaaaa caggcatgtt gaacttctca tctaaagtga
551  aagcccgcat ttatgataca aaagatatgt aa

```

This corresponds to the amino acid sequence &lt;SEQ ID 752; ORF 215.ng&gt;:

```

g215.pep
1  MKVRWRYGIA FPLILAVLGS SLSAWLGRIS EVEIEEVRNL PDEPQYTMDS
51  LDGRRFDEQG YLKEHLSAKG AKQFPENSDI HFDSPLVFF QEGRLLYEVG
101  SDEAVYHTEN KQVLFKNVNV LTKTADGRRQ AGKVETELH VDESQYAQT
151  DTPVSFQYGA SHGQAGGMTY NHKTGMLNFS SKVKAAYDT KDM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 753>:

```

m215.seq (partial)
1  ..AGCCTGTCGG CATGGTTGGG TCGTATCAGC GAAGTCGAGA TTGAAGAAGT
51  CAGGCTCAAT CCCGACGAAC CGCAATACAC AATGGACAGC TTGACCGGCA
101  GGCGGTTTGA CGAACAGGGA TACTTGAAAG AACATTGAG CGCGAAGGGC
151  GCGAAACAGT TTCCGAAAG CAGCGACATC CATTTTGATT CGCCGCATCT
201  CGTGTCTTTC CAAGAAGGCA GGTGTTGTA CGAAGTCGGC AGCGACGAAG
251  CCGTTTACCA TACCGAAAAC AAACAGGTTC TTTTAAAAA CAACGTTGTG
301  CTGACCAAAA CCGCCGACGG CAAACGGCAG GCGGGTAAAG TTGAAGCCGA
351  AAAGCTGCAC GTCGATACCG AATCTCAATA TGCCCAAAAC GATACGCCTG
401  CAGTTTCCA ATATGGTGCA TCGCACGGTC AGGCGGGCGG CATGACTTAC
451  GACCACAwA CAGGCATGTT GAACCTCTCA TCTAAAGTGA AAGCCACGAT
501  TTATGATACA AAAGATATGT AA

```

This corresponds to the amino acid sequence &lt;SEQ ID 754; ORF 215&gt;:

```

m215.pep (partial)
1  ..SLSAWLGRIS EVEIEEVRNL PDEPQYTMDS LDGRRFDEQG YLKEHLSAKG
51  AKQFPENSDI HFDSPLVFF QEGRLLYEVG SDEAVYHTEN KQVLFKNVNV
101  LTKTADGKRQ AGKVEAEKLH VDESQYAQT DTPVSFQYGA SHGQAGGMTY

```

a215	:
	YLKEHLSSKGAKQFPESSDIHFDSPHLVFFQEGRLLYEVGSDAEAVYHTENKQVLFKNNVV
	70                80                90                100              110             120
m215.pep	110          120          130          140          150          160
	LTKTADGKRQAGKVAAEKLHVDTESQYAQTDTVPVSFOYGASHGQAGGMTYDHXTGMLNFS
a215	
	LTKTADGKRQAGKVAAEKLHVDTESQYAQTDTVPVSFOYGASHGQAGGMTYDHXTGMLNFS
	130          140          150          160          170          180
m215.pep	170
	SKVKATIIYDTKDMX
a215	
	SKVKATIIYDTKDMX
	190

g216.seq (partial)

```

1      ..atgatatcga tttcgagctc ggtaccacgc gacgaaatca cgcgcatacat
51     ccccgccactc aaacgcaaag acattaccct cgtctgcac cccgcccgcc
101    ccgattcaac catggcgcg ccatgccgata tccacatcac cgcactcggt
151    tcgcaagaag cctggccggt gggggcttgc cgcaccacca gcaccaccgc
201    cgttatggct tctggcgacg cgttggcggt cgtctctctg cgcgcgccgc
251    cgttcacgcc cgacgacttc gccttgatcc accctgcccg cagcctcggc
301    aaacgcctgc ttttgcgcgt tgcgcacatt atgcacaaag gcggcggcct
351    gcccgccgct cgaactcgga cgcctctgaa aggagccatc gtcagcatga
401    gcgagaaaag tttgggcata tgggcgggaa cggacgggca aaggctgtct
451    qaagaacqctt ttacttqa

```

g216.pep (partial)

```

1      .MISISSVPS DEITAIIPAL KRKIDITLVCIT TARPDSMAR HADIHITASV
51     SQBACPGLA PTTSTAVMA LGDALAVVLL RARAFTPDDF ALIHPAGSLG
101    KRLLLLRADI MHKGGGLPAV RLGTPLKGAI VSMSEKGLGM WAGTDGQRLS
151    ERPFY*

```

m216.seq

1	ATGGCAATGG	CAGAAAAACGG	AAAATATCTC	GACTGGGCAC	GCGAAGTGT
51	GCACGCCGAA	GCGGAAGGCT	TGCGCGAAAT	TGCAGCGGAA	TTGACAAAA
101	ACTTCGTCTT	TGCGGCAGAC	GCSTGTATTG	ACTGCAAGG	CAGGGTCGTT
151	ATCACGGCTA	TGGTCAAGTC	GGGACATATC	GGGCGCAAAA	TGGCGGCAAC
201	TATGCGCTCG	ACCGGCACGC	CTGCGTTTTT	CGTCACCCCT	CGGGAAGCGG
251	CACACGCGA	TTTGGGTATG	ATTGTGGACA	rCGACGTGGT	CGTCGCGATT
301	TCCAATTCCG	GCGAAAGCGA	CGAAATCGCC	GCCATCATCC	CGCAGCTCAA
351	ACGCAAAAGC	ATCACGCTTG	TCTGCATCAC	CGCCCCCCCC	GATTCAACCA
401	TGGCGCGCCA	TGCCGACATC	CACATCACGG	CGTCGGTTTC	CAATGAAGCC
451	TGCCCGCTGG	GGCTTGCCCC	GACCACCAGC	ACCACCGCGC	TCATGGCTTT
501	GGGCGATGCG	TTGGCGGTCT	TCTcGTCTGG	CgcACGCGCG	TTACGCCCCG
551	ACGATTTCGC	CTTGAGCCAT	CCTGCGGCA	GCCTCGGCAA	ACGCTACTTT
601	TTGCGCGTTG	CCGACATTAT	GCACAAAAGC	GCGCGCCTGC	CTGCCGTCGG
651	ACTCGGCACG	CCCTTGAAG	AAGCCATCGT	CAGCATAGT	GA AAAAGGGC
701	TGGGATGTTT	GGCGGTAAAC	GACGGGCAAG	CGCGTCTGAA	AGCGGTATTCT
751	ACCGACGGCG	ATTTGCGCCG	CCTGTTTCAA	GAATGCGACA	ATTTTACCGG
801	TCTTTTCGATA	GACGAAGTCA	TGCATACGCA	TCCTAA AACC	ATCTCCGCCG
851	AACGTCTCGC	CACCGAAGCC	TCGAAAGTCA	TGCAGGCAAA	CCATGTGAAC
901	GGGCTTCTGG	TTACCGATGC	AGATGGCGTG	CTGATCGGCG	CGCTGAATAT
951	GCACGACCTG	CTGGCGGCAC	GGATTGTATA	G	

m216.pep

1 MAMAENKYL DWAREVLHAE AEGLREIAAE LXKNFVLAD ALLHCKGRVV

301 GLLVTDADGV LIGALNMHDL LAARIV\*

m216/a216 97.2% identity in 326 aa overlap

	10	20	30	40	50	60
m216.pep	MAMAENKYL	DWAREVLHAE	AEGLREIAAE	LXKNFVLAAD	ALLHCKGRVV	ITGMVKS
a216	MAMAGNEKYL	DWAREVLHTE	AEGLREIAAD	LDENFALAAD	ALLHCKGRVV	ITGMGKS
	10	20	30	40	50	60
	70	80	90	100	110	120
m216.pep	GRKMAATMA	STGTPAFFV	HPAEAAHGD	LGMIVDXDV	VVAISNSGE	SDEIAAII
a216	GRKMAATMA	STGTPAFFV	HPAEAAHGD	LGMIVDNDV	VVAISNSGE	SDEIAAII
	70	80	90	100	110	120
	130	140	150	160	170	180
m216.pep	ITLVCITAR	PDSTMARHAD	ITITASVSKE	ACPLGLAPT	TSTTAVMAL	GDALAVVLL
a216	ITLVCITAR	PDSTMARHAD	ITITASVSKE	ACPLGLAPT	TSTTAVMAL	GDALAVVLL
	130	140	150	160	170	180
	190	200	210	220	230	240
m216.pep	FTPDDFALS	HPAGSLGKR	LLLRVADIM	HKGGGLPAV	RLGTPPLKE	AIIVSMSEK
a216	FTPDDFALS	HPAGSLGKR	LLLRVADIM	HKGGGLPAV	RLGTPPLKE	AIIVSMSEK
	190	200	210	220	230	240
	250	260	270	280	290	300
m216.pep	DGQGRCLK	GVFTDGD	LRLRFQEC	NFTGLSID	EVMTHTPK	TISAERLATE
a216	DGQGRCLK	GVFTDGD	LRLRFQEC	NFTGLSID	EVMTHTPK	TISAERLATE
	250	260	270	280	290	300
	310	320				
m216.pep	GLLVTDADG	VLIGALNMH	DLLAARIVX			
a216	GLLVTDADG	VLIGALNMH	DLLAARIVX			
	310	320				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 763>:

g217.seq

```

1  atggcggatg acggtttgtt gcggcaactg tccgaaaaac ccagccaaag
51  tgctctcttc ctgccatttg acccattcgt ttctgaggtt ttggactgcc
101 ttttggatcat cgggcccggc ttgaacaat gtttcaagca aatcccggca
151 acgcgccacc cattcgccga cgtcgccagg ttgccgccat atccgggcaa
201 tatccgacag ggtttcgagg aaggcggcaa aacgtccgaa catggcgggt
251 tgattcacgt cggcatacca cgcgctgaca tcctgccaca tcgggttgcc
301 gccttcgggc agcatccagc ccaatatcat acggtctgcc gcctgcttcc
351 aggtaaacag ctgatccgtg ccgccgcgca tttctccgtc caatcccaaa
401 tggacgttca aatcggcaac catatcgtgc aaaagcggca aatcgtcccc
451 ggtcagtcgc aaacggcgca acacggggcg gggttccaaa agcgcgagca
501 ctttgccgac ttcaaaacgg ctttccagca agtcggacac gcactccaac
551 gcataaaaaa acggttgccg gcggtgatt ttacagtcgc aaacggaata
601 cggcaatgcc tgcgcgccgg gttgcgcctg tccgaacacg gcttcataa
651 aagcgtata gggttcgata ttcgggggta a

```

This corresponds to the amino acid sequence <SEQ ID 764; ORF 217.ng>:

g217.pep..

```

1  MADDGLLRQL SEKPSQSALF LPFDPFVFEV LDCLLVIGPG LKQCFKQIPA
51  TRHPFADRRR LPPYPGNIRQ GFEEGGKTSE HGGLIHVGIP RADILPHRVA
101 AFGQHPAQYH TVCRLLPGKQ LIRAAHFSV QSPMDVQIGN HIVQKRQIVP
151 GQSETAQHGR GFQKREHFAD FKTAFOQVGH ALQRIKKRLP AADFHVRNGI

```

501

```

201 TATCCGTCAG GGTTCGAGG AAGGCGGCAA AACGTCCGAA CAGGCGGGTT
251 TGGTTCACGT CGGCATACCA CGCGCTGACC CCCTGCCACA TCGGATTGCC
301 GCCTTCGGGC AGCATCCAGC CCAATATCAT GCGTCTACC GCCTGCTTCC
351 AGGTGAACAG CTGATCCGTG CCGCCGCGCA TTTCTCCGTC CAAACCCAG
401 CGGACGTTC AATCGGCAAC CATGTCGTGC AAAAGCGGCA AATCGTCCTC
451 AGTCAGTCCG AAATGGCGCA ACACGGGCGC GGTTCCTAAA AGCACAAAGCA
501 CTTTATCGAC TTCAAATCGG CTTTCCAACA AGTCGAACAG GCATGACAAA
551 GCATGAAACA GCGGTTGTCG GCGGCTGATT TTCACATCCG AAACGGAATA
601 CGGCAATGCC TGCGCGCCCG GCTGCGCCTG TCCGAACACG GCTTCGATAA
651 AAGGCGTATA GGATTCGATA TTCGGGGTTA A

```

This corresponds to the amino acid sequence <SEQ ID 768; ORF 217.a>:

```

a217.pep
1 VADDGVQRQL SGKLRQFGFR LPFDPFVFEEA LDCLLVIAFD LEQCFKQIPA
51 TRHPPVNNRR LPPYPYNIRO GFEEGGKTSE QGGLVHVGI RADPLPHRIA
101 AFGQHPAQYH AFYRLLPGEQ LIRAAAHFSV QTPADVQIGN HVVQKRQIVL
151 SQSEMAQHGR GF*KHKHFID FKSAFQQVEQ A*QSMKQRLS AADFHIRNGI
201 RQCLRAGLRL SEHGFDKRR I GFDIRG*

```

m217/a217 90.3% identity in 226 aa overlap

	10	20	30	40	50	60
m217.pep	MADDGVRRQLSGKLRQFGFRLPFDPFVFKVLDXLLVIGFSLEQCFKQIPATRHPPFADRCG					
a217	VADDGVQRQLSGKLRQFGFRLPFDPFVFEEALDCLLVIAFDLEQCFKQIPATRHPPVNNRR					
	10	20	30	40	50	60
m217.pep	LPPYPYNIROGFEEGGKTSEHGGLIHVGI PRADILPHRIA AFGQHPAQYHAFYRLLPGEQ					
a217	LPPYPYNIROGFEEGGKTSEQGGLVHVGI PRADPLPHRIA AFGQHPAQYHAFYRLLPGEQ					
	70	80	90	100	110	120
m217.pep	LIRAAAHFSVQTPVDVQIGNHVQKRXIVLSQSETAQHGRGFXKHKHFIDFKSAFQQVEQ					
a217	LIRAAAHFSVQTPADVQIGNHVQKRQIVLSQSEMAQHGRGFXKHKHFIDFKSAFQQVEQ					
	130	140	150	160	170	180
m217.pep	AXQSMKQRLSAAADFHVXHGIRQCLRTGLRLSEHGFDKRRIGFDIRGX					
a217	AXQSMKQRLSAAADFHIRNGIRQCLRAGLRLSEHGFDKRRIGFDIRGX					
	190	200	210	220		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 769>:

```

g218.seq
1 atggttgccg tggatcctta tacggcaaaa gtggtcaaca ccatgccgcg
51 caatcagggg tggatcaca ctatggatga aatccacggc gatatgatgc
101 tcggtgcggc aggcgattat cttttggaaa cggcagcttc actgaccatt
151 attatggttg tcagcggttc gtacctttgg tggcgcaaac agcgcggcat
201 taaagcgatg ctgctgccgc caaaaagcag ggcgcgttct tgggtggcga
251 atctgcacgg cgcgttttga acttgggtgt cgttgatttt actgtgttct
301 tgccctgtcg gtattgcttg ggcaggtatt tggggcgcca aattcgtgca
351 ggcttggaat cagttcccgc ccggcaaatg ggtgtcgaa ccgaaccccg
401 tttcaatcgt gccgaccac gccgaggtat tgaatgacgg caagggttaag
451 gaagtgccgt ggatttttga gcttatgcct atgctgtct cagggacgac
501 tgtgggtgaa aacggcatta accccaccga gcccaataac attggaacc
551 gtcgaccgtt tcgcgcggga aatcggttcc aaaggcggt atcagttgaa
601 tttgcccaaa ggcgaggacg gggatatggac tttgcgcag gattctatga
651 gttatga

```

This corresponds to the amino acid sequence <SEQ ID 770; ORF 218.ng>:

```

g218.pep
1 MVAVDPTYAK VVNTMPRNQG WYHTMDEIHG DMMLGAAGDY LEETAASLTI

```

503

**a218.seq**

```

1  ATGGTCGCGG TCGATCCTTA TACGGCAAAA GTGGTCAGTA CCATGCCGCG
51  CAATCAGGGT TGGTATTACG CGATGGATGA AATCCACAGC GATATGATGC
101 TCGGTTCGAC AGGTGATTAT CTTTGGAAA CCGCTGCATC GCTGACGATT
151 ATCATGATAA TCAGCGGTTT GTACCTTTGG TGGGTGAAAC GGCGCGGCAT
201 CAAGGCGATG CTGCTGCCGC CAAAAGGCAG GGCGCGTTCT TGGTGGCGGA
251 ATCTGCACGG CGCGTTTGA ACTTGGGTGT CGTTGATTTT ACTGTTGTTT
301 TGCCTGTCGG GTATTGCTTG GGCAGGTATT TGGGGCGGCA AGTTCTGTGA
351 GGCTTGGAGT CAGTTCCCGG CAGGCAAATG GGGTGTGAA CCGAACCCCTG
401 TTTCAGTCGT GCCGACCCAC GGCGAGGTAT TGAATGACGG CAAGGTTAAG
451 GAAGTGCCGT GGGTTTGA GCTTACGCCT ATGCCTGTTT CAGGGACGAC
501 TGTGGGCAAA GACGGTATTA ACCCTGACGA GCCGATGACA TTGGAACCG
551 TCGACCGTTT TGC GCGG.GA AATCGGTTT AAAGGCGGTT ATCAGCTGAA
601 TTTGCCCAAA GGCGAGGACG GCGTATGGAC TTTGTCGCAG GATTCTATGA
651 GTTA

```

This corresponds to the amino acid sequence <SEQ ID 774; ORF 218.a>:

**a218.pep**

```

1  MVAVDPYTAK VVSTMPRNQG WYYAMDEIHS DMMLGSTGDY LLETAASLTI
51  IMIISGLYLW WVKRRGIKAM LLPPKGRARS WWRNLHGAFG TWVSLILLLF
101 CLSGIAWAGI WGGKFVQAWS QFPAGKWGVE PNPVSVVPTH GEVLNDGKVK
151 EVFWVLELTP MPVSGTTVGK DGINPDEPMT LETVDRFARX NRFQRALSAE
201 FAQRRGRMD FVAGFYEL

```

**m218/a218 95.9% identity in 218 aa overlap**

	10	20	30	40	50	60
m218.pep	MVAVDPYTAKVVSTMPRNQGWYYTMDEIHS	DMMLGAAGDYLLETAASLTI	IMVVSGLYLW			
a218	MVAVDPYTAKVVSTMPRNQGWYYAMDEIHS	DMMLGSTGDYLLETAASLTI	IMIISGLYLW			
	70	80	90	100	110	120
m218.pep	WVKRRGIKAMLLPSKXARSWWRNLHGT	FGTWVSLILLFCLSGIAWAGIWGGKFVQAWS				
a218	WVKRRGIKAMLLPPKGRARSWWRNLHGAF	GTWVSLILLFCLSGIAWAGIWGGKFVQAWS				
	130	140	150	160	170	180
m218.pep	QFPAGKWGVEPNPVSVVPTHGEVLNDGK	VEVPWVLELTPMPVSGTTVGKDGINPDEPMT				
a218	QFPAGKWGVEPNPVSVVPTHGEVLNDGK	VEVPWVLELTPMPVSGTTVGKDGINPDEPMT				
	190	200	210			
m218.pep	LETVDRFARXNRFQRALSV	FAQRRGRMDFVAGFYEL				
a218	LETVDRFARXNRFQRALSAE	FAQRRGRMDFVAGFYEL				

505

	70	80	90	100	110	120
	130	140	150	160	170	180
m219.pep	LGWWSVLANVLFCLAVIFIGISGCVMMWKRRPTGAVGIVPPAQKVKLPVWMMALPLLAI					
	:     :     :     :     :					
g219	LGWWSVLANVVFLAVIFIGISGCVMMWKRRPSGVAGIVPPAQKIKLPVWAMALPLLLI					
	130	140	150	160	170	180
	190	200	210			
m219.pep	ALLFPTSLLAIAVIWLLDTLLSRIPVLRWFKX					
	:     :     :     :					
g219	ALLFPTALLAIAVIWLLDTLLSRIPVLRWFK					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 779>:

```

a219.seq
1  ATGACGGCAA GGTAAAGGAA GTGCCGTGGG TTTTGGAGCT TACGCCTATG
51  CCTGTTTCAG GGACGACTGT GGGCAAAGAC GGTATTAACC CTGACGAGCC
101 GATGACATTG GAAACCGTCG ACCGTTTTCG GCGG.GAAAT CGGTTTCAAA
151 GGGCGTTATC AGCTGAATTT GCCCAAAGGC GAGGACGGCG TATGGACTTT
201 GTCGCAGGAT TCTATGAGTT ACGACATCAT CAGCCCGTTT GCTGACCGCA
251 CGGTGATATC CGACCACTAC AGCGGCAAGA TTCTTGCCGA CATCCGTTTT
301 GACGATTACA ACCCGTTCGG CAAATTTATG GCGGAGCA TTGCGCTGCA
351 TATGGGGACT TTGGGCTGGT GGAGCGTGTG GGCCAAUGTT TTGTTCTGCC
401 TTGCCGTGAT TTTTATCGGC ATCAGCGGCT GCGTGATGTG GTGGAAACGC
451 CGTCCGTCCG GCGCGTGGG CATGGTCCG CCGGCGCAAA AAATCAAGCT
501 GCCCCGTCTG TGGGCAATGG CCGTGCCGCT GCTGCTGATT GCATTGCTTT
551 TCCCCACCGC GTTGCTTGCC ATTGCCGTGA TTTGGCTGTT GGATACGCTG
601 CTGTTGTTCG GGATTCCTGT TTTGAGGAGA TGGTTTAAAT GA

```

This corresponds to the amino acid sequence <SEQ ID 780; ORF 219.a>:

```

a219.pep
1  MTARLRKCRG FWSLRRLCLFQ GRLWAKTVLT LTSR*HWKPS TVLRXEIGFK
51  GRYQLNLPKG EDGVWTLSD SMSYDMISPF ADRTVHIDQY SGKILADIRF
101 DDYNPFGKFM AASIALHMG T LGWWSVLANV LFCLAVIFIG ISGCVMMWKR
151 RPSGAVGMVP PAQKIKLPVW WAMAVPLLLI ALLFPTALLA IAVIWLLDTL
201 LLSRIPVLRW WFK*

```

m219/a219 94.8% identity in 213 aa overlap

	10	20	30	40	50	60
m219.pep	MTARLRKCRGFWSLRRLCLFQGRXWAKTALTLSRXHWKPSTALRGEIGFKGRYQLNLPKG					
	:     :     :     :     :					
a219	MTARLRKCRGFWSLRRLCLFQGRWLWAKTVLTLSRXHWKPSTVLRXEIGFKGRYQLNLPKG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m219.pep	EDGVWTLSDSMSYDMISPFADRTVHIDQYSGKILADIRFDDYNPFGKFMAASIALHMG					
	:     :     :     :     :					
a219	EDGVWTLSDSMSYDMISPFADRTVHIDQYSGKILADIRFDDYNPFGKFMAASIALHMG					
	70	80	90	100	110	120
	130	140	150	160	170	180
m219.pep	LGWWSVLANVLFCLAVIFIGISGCVMMWKRRPTGAVGIVPPAQKVKLPVWMMALPLLAI					
	:     :     :     :     :					
a219	LGWWSVLANVLFCLAVIFIGISGCVMMWKRRPSGAVGMVPPAQKIKLPVWAMAVPLLLI					
	130	140	150	160	170	180
	190	200	210			
m219.pep	ALLFPTSLLAIAVIWLLDTLLSRIPVLRWFKX					
	:     :     :     :					
a219	ALLFPTALLAIAVIWLLDTLLSRIPVLRWFKX					
	190	200	210			

```
a221.seq
1   ATGGTGGTGT   TGATGCTCCG   AAGTCTGGTG   CGGCAGGCCG   TAAATCAAA
51  CGACGCTGAC   GGATTTTGAAC   CCCGCTTCGC   GCGCGCGCAT   GATGACTTCT
101 TTGGTTTCTT   CGTAACTTTG   GATGCGGTGG   ACCGCGCGCT   GCACCTTTGG
151 GTCGAAATCC   TGAATGCCGA   TGCTCATGCG   GTTGAAGCCG   AGTCTGCCGA
201 GCATGAGGAC   GGTGTCCGCG   CTGACTTTGC   GCGGGTCGAT   TTCGATGGAG
251 TATTCGCCGG   TGGGGATTAA   CTCGAAATGT   TTGCGTATCA   TCGCGGAAGC
301 ACGTTCGATT   TGGTCGTGCG   TCAAAAAGGT   CGGCGTCGCG   CGCCCGAAGT
351 GCAGTTGGGC   AAGCTGGTGC   CGTCCGTTCA   GATGTGGAGC   GAGCAGTTCC
401 ATTTCITTTT   CAAGAATTG   CGTGTAGGCA   TCGCGCGCGC   TTTTGTCTTT
451 GGTGATGATT   TTGTTGCAGC   GCGCAGTAGA   GCAGATGGTG   TTGCAGAACG
501 GAATGTGAAT   GTAAGGGGAA   AGCGGTTTGT   TTA
```

a221.pap

1 MVVLMRLSLV RQAVNQIDAD GFEPFRFARRI DDFGFFFTVL DAVDRRLHEFG  
51 VEILNADAHA VEAESAIEDH GVAADFARVD FDGVFAGGD\* LEMFAYHAED  
101 TFDLVVAQKG RRAAAEVQLG KLVPSVQMVS EQFHFFFKKF DVGIGAAEFV  
151 GDDEFAAAVV ADGVAERNVN VVGKRFV\*

	10	20	30	40	50	60
m221.pep	MXVLMXRSILVRQAVNQIDADGFEPFRFARRIDDFGGFVTLD	DAVDRLHFGVEILNADAH				
a221	MVVLMLRSILVRQAVNQIDADGFEPFRFARRIDDFGGFVTLD	DAVDRLHFGVEILNADAH				
	10	20	30	40	50	60
	70	80	90	100	110	119
m221.pep	VEAESAEHEDGVAADFARVDFDGVFAGGD	XL	EMFAYHAEDT	FDLFVAQKGA	-CPAEVQLG	
a221	VEAESAEHEDGVAADFARVDFDGVFAGGD	XL	EMFAYHAEDT	FDLFVAQKGRRAAEVQLG		
	70	80	90	100	110	120
	120	130	140	150	160	170
m221.pep	KLVP	SVQMWSEQHFHFFFKI	FDV	GIGAA	FVFGDDFVAAAVVADGVAERNVN	VVKGRFVX
a221	KLVP	SVQMWSEQHFHFFFKI	FDV	GIGAA	FVFGDDFVAAAVVADGVAERNVN	VVKGRFVX
	130	140	150	160	170	

g223.seq

```
1 atggaattca ggcaccaggt agtggtagtt ggtgtcgaac catttggtca
51 tttcgatggc gaattggtct ttgttgccgc gcgccagttg gaagaattgt
101 tccaaaggca ggttttggct atcgaaggcg aaacgggcgg gaatcgccgc
151 cgtggatact tcgaggtcga ggatgtgatg gtagaagaatg aaatcacgta
201 cagcaacgta atcagcgcta ggagcagctt ggtgtttcca gttttttctg
251 ccgaggtctt tggcaacgtc gagcagctct tgttcaetga tctcttttgcg
301 ccagtatatt tcttgggcga atttcaattc acggaaggcg ccgacacgcg
351 qaaagcctga
```

q223.ppt...

70 80 90 100 110 120

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 793>:

```
g225.seq
1 atggattctt ttttcaaacc ggcagtttgg gcggttttgt ggctgatgtt
51 tgcggtccgc cccgcccttg ccgacgagtt gaccaacctg ctcagcagcc
101 gcgagcagat tctcagacag tttgccgaag acgaacagcc cgttttacc
151 gtcaaccgag ccccgcccg gcggggcggc aatgccgacg aactcatcgg
201 cggcgcgatg gggcttaacg aacagcccgt tgtacgcgtc aaccgagccn
251 ccgcccggcg gcggggcaat gccgacaaac tcacggcgag cgcgatgcgg
301 cttttgggta ttgcctaccg ctacggcggc acatgggtgt ctaccggttt
351 tgactgcagc ggatttatgc agcacatctt caaacgcgcc atgggcatca
401 acctgccgcg cagtcggcg gaacaggcgc ggatgggcgc acccgttgcc
451 cgaagcgaat tgcagcccgg ggatatggtg tttttccgca cgcctggcgg
501 cagccgcatt tcccatgtcg gactttatat cggcaacaac cgtttcatcc
551 acgcgccgcg cagggggaaa aatatcgaaa tcaccagcct gagccacaaa
601 tattggagcg gcaaatatgc gttcggccgc cgggtcaaga aaaacgaccc
651 gtcacgcttt ctgaactga
```

This corresponds to the amino acid sequence <SEQ ID 794; ORF 225.ng>:

```
g225.pep
1 MDSFFKPAVW AVLWLMFAVR PALADELTNL LSSREQILRQ FAEDEQPVLP
51 VNRAPARRAG NADELIGGAM GLNEQPVVRV NRAXARRAGN ADKLIGSAMR
101 LLGIAYRYGG TSVSTGFDCS GFMQHIFKRA MGINLPRTSA EQARMGAPVA
151 RSELQPGDMV FFRTLGGSR I SHVGLYIGNN RFIHAPRTGK NIEITSLSHK
201 YWSGKYAFAR RVKKNPSRF LN*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 795>:

```
m225.seq (partial)
1 ..TTTCAAACC CGGCAGTTTG GCGGTTTTG TGGCTGAWGT TTGCCGTCCG
51 CCCC GCCCTT GCCGACGAGT TGACCAACCT GCTCAGCAGC CGCGAGCAGA
101 TTCTCAGACA GTTTGCCGAA GACGAACAGC CCGTTTTACC CATCAACCGA
151 GCCCCGCCCG GCGGGCGGG CAATGCCGAC GAACTCATCG GCAGCGCGAT
201 GGGGCTTAAC GAACAGCCCG TTTTACCCGT CAACCGAGTC CCGCCCCGGC
251 GGGCGGGCAA TGCCGACGAA CTCATCGGCA ACGCGATGGG GCTTAACGAA
301 CAGCCCCGTTT TACCCGTCAA CCGAGCCCCC GCCCGGCGGG CGGGCAATGC
351 CGACGAACTC ATCGGCAACG CGATGGGACT TTTGGGTATT GCCTACCGCT
401 ACGGCGGCAC ATCGGTTTCT ACCGGTTTTG ACTGCAGCGG CTTCATGCAG
451 CACATCTTCA AACCGGCCAT GGGCATCAAC CTGCCGCGCA CGTCGGCAGA
501 ACAGGCACGG ATGGGTACGC CGGTTGCCCG AAGCGAATTG CAGCCCCGAG
551 ATATGGTGTT TTTCCGCACG CTCGGCGGCA GCCGCATTTC CCATGTCGGA
601 CTTTATATCG GCAACAACCG CTTTCATCCAC GCGCCGCGCA CGGGGAAAAA
651 TATCGAAATC ACCAGCCTGA GCCACAAATA TTGGAGCGGC AAATACGCGT
701 TCGCCCGCGG GGTCAAGAAA AACGACCCGT CCCGCTTTCT GAACTGA
```

This corresponds to the amino acid sequence <SEQ ID 796; ORF 225>:

```
m225.pep (partial)
1 ..FSNPAVWAVL WLXFAVRPAL ADELTNLLSS REQILRQFAE DEQVLPINR
51 APARRAGNAD ELIGSAMGLN EQVLPVNRV PARRAGNADE LIGNAMGLNE
101 QVLPVNRAP ARRAGNADEL IGNAMLLGI AYRYGTSVS TGFDCSGFMQ
151 HIFKRAMGIN LPRTSAEQAR MGTTPVARSEL QPGDMVFFRT LGGSRISHVG
201 LYIGNNRFIH APRTGKNIEI TSLSHKYWSG KYAFARRVKK NDPSRFLN*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 225 shows 83.5% identity over a 248 aa overlap with a predicted ORF (ORF 225.ng) from *N. gonorrhoeae*:

m225/g225

```

10      20      30      40      50
m225.pep  FSNPAVWAVLWLXFAVRPALADELTNLLSSREQILRQFAEDEQVLPINRAPARRAG
          | : ||||| | | | | | | | | | | | | | | | | | | | | | | | |
g225      MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQVLPVNRAPARRAG
```



511

```

a225      |||||||
          NADELIGSAMGLNEQPVLPVNRXPARRAGNADXLIGNAMGLNEQPVLPVNRVPARRAGNA
          70      80      90      100     110     120

m225.pep  90      100     110     120     130     140
          DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
          |||||||
a225      DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF
          130     140     150     160     170     180

m225.pep  150     160     170     180     190     200
          MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
          |||||||
a225      MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVXFFRTLGGSRISHVGLYIGNNRF
          190     200     210     220     230     240

m225.pep  210     220     230     240     249
          IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDFSRFLNX
          |||||||
a225      IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDFSRFLNX
          250     260     270     280

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 799>:

g225-1.seq

```

1   atggattctt ttttcaaacc ggcagtttgg gcggttttgt ggctgatgtt
51  tgcgctccgc cccgcccttg ccgacgagtt gaccaacctg ctcagcagcc
101 gcgagcagat tctcagacag ttgcccgaag acgaacagcc cgttttacct
151 gtcaaccgag ccccgcccg gggggcgggc aatgccgacg aactcatcgg
201 cggcgcgatg gggcctaacg aacagcccg tgtacgcgtc aaccgagccn
251 ccgcccgcg ggcgggcaat gccgacaaac tcacgcgcag cgcgatcgcg
301 cttttgggta ttgcctaccg ctacggcggc acatcggtgt ctaccggttt
351 tgactgcagc ggattcatgc agcacatctt caaacgcgcc atgggcatca
401 acctgccgcy cagtcggcg gaacaggcgc ggatggcgcc acccgttgcc
451 cgaagcgaat tgcagcccg ggatatgggt ttttccgca cgcctcgcg
501 cagccgcatt tcccatgtcg gactttatat cggcaacaac cgttcctacc
551 acgcgcgcgc cacggggaaa aatatcgaaa tcaccagcct gagccacaaa
601 tattggagcg gcaaatatgc gtccgcccgc cgggtcaaga aaaacgaccc
651 gtcacgcttt ctgaactga

```

This corresponds to the amino acid sequence <SEQ ID 800; ORF 225-1.ng>:

g225-1.pep

```

1   MDSFFKPAV AVLWLMFAVR PALADELTNL LSSREQILRQ FAEDEQPVLP
51  VNRAAPARRAG NADELIGGAM GLNEQPVVRV NRAXARRAGN ADKLIGSAMR
101 LLGIAYRYGG TSVSTGFDCS GFMQHIFKRA MGINLPRTSA EQARMGAPVA
151 RSELQPGDMV FFFRTLGGSR SHVGLYIGNN RFIHAPRTGK NIEITSLSHK
201 YWSGKYAFAR RVKKNDFSRF LN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 801>:

m225-1.seq

```

1   ATGGATTCTT TTTTCAAACC GGCAGTTTGG GCGGTTTGTG GGCTGATGTT
51  TGCCGTCCGC CCCGCCCTTG CCGACGAGTT GACCAACyTG CTCAGCAGCC
101 CGGAGCAGAT TCTCAGACAG TTTGCCGAAG ACGAACAGCC CGTTTACCC
151 ATCAACCGAG CCCC GCCCG GCGGGCGGGC AATGCCGACG AACTCATCGG
201 CAGCGCGATG GGGCTTAACG AACAGCCCGT TTTACCCGTC AACCGAGTCC
251 CCGCCCGGCG GCGGGGCAAT GCCGACGAAC TCATCGGCAA CGCGATGGGG
301 CTTAACGAAC AGCCCGTTT ACCCGTCAAC CGAGCCCCCG CCGCGCGGCG
351 GGGCAATGCC GACGAACTCA TCGGCAACGC GATGGGACTT TTGGGTATTG
401 CCTACCGCTA CGGCGGCACA TCGGTTTCTA CCGGTTTGA CTGCAGCGGC
451 TTCATGCAGC ACATCTTCAA ACGCGCCATG GGCATCAACC TGCCGCGCAC
501 GTCGGCAGAA CAGGCACGGA TGGGTACGCC GGTGCCCCGA AGCGAATTGC
551 AGCCCCGAGA TATGGTGT TTCCGCACGC TCGGCGGCAG CCGCATTTCC
601 CATGTCGGAC TTTATATCGG CAACAACCGC TTCATCCACG CGCGCGCAC
651 GGGGAAAAAT ATCGAAATCA CCAGCCTGAG CCACAAATAT TGGAGCGGCA
701 AATACCGGTT CGCCCGCCGG GTCAAGAAAA ACGACCCGTC CCGCTTCTG
751 AACTGA

```

513

a225-1/m225-1 88.6% identity in 280 aa overlap

```

      10      20      30      40      50      60
a225-1.pep MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRXPARRAG
            |||
m225-1      MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
            |||

      70      80      90     100     110     120
a225-1.pep NADELIGSAMGLNEQPVLPVNRXPARRAGNADXLIGNAMGLNEQPVLPVNRVPARRAGNA
            |||
m225-1      NADELIGSAMGLNEQ-----VLPVNRVPARRAGNA
            |||

     130     140     150     160     170     180
a225-1.pep DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF
            |||
m225-1      DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
            |||

     190     200     210     220     230     240
a225-1.pep MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVXFRTLGGSRIHVGLYIGNNRF
            |||
m225-1      MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRIHVGLYIGNNRF
            |||

     250     260     270     280
a225-1.pep IHAPRTGKNIETSLSHKYWSGKYAFARRVKKNPDRFLNX
            |||
m225-1      IHAPRTGKNIETSLSHKYWSGKYAFARRVKKNPDRFLNX
            |||

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 805>:

```

g226.seq
1  ATGAGCGAAA TCCTCAGGCA GCCCAGCGTT CTGCTTTTCC TCACGCTTGC
51  CGTGTACGCG CTTGCGATTA TCGTGCGCAC GCGCACGGGC AATATCTTCT
101 GCAACCCCGT ACTCGTCAGC ACTATCGTGC TGATTGCCTA CCTGAAAATC
151 CTCGGTATCG ATTATGCGGT GTACCACAAC GCCGCGCAAT TCATTGATTT
201 TCGGCTGAAA cccGccgtCG TCGTGCTTGC CGTGCCGCTC TACCAAAACC
251 GCCGTAAAT CTTCACCAG TGGCTGCCCG TCATCGTTTC GCAGCTTGCG
301 GGCAGCGTTA cggGCATTGT tacggggATG TATTTTgccc ctgggctcgg
351 gccggatacc caattctcct tcccgccctcg tcttcaatat ctgttattta
401 caccctctgg aatcccaatt cacaccctgt atgcgcgggt tctcccgcca
451 tttctgttgc ctccgcctct cctgccgcgc ctccgcccgc atacattgcg
501 ccggttcaca atacttccaa aaaaactacg gccgtttaag cccctcctcc
551 cagttgtggt cctttctcct Ccgggcctcg cccctcccct cttataa

```

This corresponds to the amino acid sequence <SEQ ID 806; ORF 226.ng>:

```

g226.pep
1  MSEILRQPSV LLFLTLAVYA LAIIVRTRTG NIFCNPVLVS TIVLIAYLKI
51  LGIDYAVYHN AAQFIDFRLK PAVVVLAVPL YQNRKIFNQ WLPVIVSQLA
101 GSVTGIVTGM YFAAWLGPDT QFSFPPRLQY LLFTPSGIPI HTLYARVLP
151 FLLPPPLPR LGPHTLRRFT ILPKKLRFK PLLPVVLSP PGLAPPLL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 807>:

```

m226.seq
1  ATGAACGAAA TCCTCAGGCA GCCCAGCGTT CTGCTTTTCC TCACGCTTGC
51  CGTGTACGCG CTTGCGATTA TCGTGCGCAC GCGCACGGGC AATATCTTCT
101 GCAACCCCGT ACTCGTCAGC ACTATCGTGC TGATTGCCTA CCTGAAAATC
151 CTCGGTATCG ATTATGCGGT GTACCACAAC GCCGCGCAAT TCATTGATTT
201 TTGGCTGAAA CCCGCCGTCG TCGTGCTTGC CGTGCCGCTC TACCAAAACC
251 GCCGTAAAT CTTCACCAG TGGCTGCCCG TCATCGTTTC ACAGCTTGCG
301 GGCAGCGTTA CGGGCATTGT TACAGGGATG TATTTTGCCA AATGGCTGGG
351 CGCGGAACGC GAAGTCGTCC TCTCGCTCGC GTCCAAATCT GTTACCAACC
401 CCATCGCTAT TGAAATCACC CGCTCCATCG GCGGCATTCC CGCCATTACC
451 GCCGCCACCG TCATCATTGC CGGTCTGGTC GGACAGATTG CCGGTTACAA

```

515

```

m226.pep      MNEILRQPSVLLFLTLAVYALAIIVRTRTGNIFCNPLVSTIVLIAYLKILGIDYAVYHN
a226          MNEILRQPSILLFLTLAVYALAIIVRTRTGNIFCNPLVSTIVLIAYLKILGIDYAVYHN
              10      20      30      40      50      60

              70      80      90      100     110     120
m226.pep      AAQFIDFWLKPAVVVLAVPLYQNRKIFNQWLPVIVSQLAGSVTGIVTGMYFAKWLGAER
a226          AAQFIDFWLKPAVVVLAVPLYQNRKIFNQWLPVIVSQLAGSVTGIVTGMYFAKWLGAER
              70      80      90      100     110     120

              130     140     150     160     170     180
m226.pep      EVVLSLASKSVTNPIAIEITRSIGGIPAITAATVIIAGLVGQIAGYKMLKNTVMPSSVG
a226          EVVLSLASKSVTNPIAIEITRSIGGIPAITAATVIIAGLVGQIAGYKMLKNTVMPSSVG
              130     140     150     160     170     180

              190     200     210     220     230
m226.pep      MSLGTASHAMGIAASLERSRMAAYAGLGLTFNGVLTALIAPLLPVLGFX
a226          MSLGTASHAMGIAASLERSRMAAYAGLGLTFNGVLTALIAPLLPVLGFX
              190     200     210     220     230

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 811>:

```

g227.seq
1  atgaacatca tccgcgcgct cctcatcate ctgcgctgcc tcgcccgcgg
51  cgaaaccgcc gttttcctag caggcatcaa actgcccggc agcatcgctc
101 gcatggcgct gctgtttgcg cttttgcagg cgggttggtc caaaacgtct
151 tggctgcaac agcttaccga cgcgctgatg gcaaaccctga cgctgttctt
201 cgtgccgccc tgcgtggcgg tcatcagcta ttggatttg attgccgacg
251 attggtttcc gatactgggt tccgcctccg ccagcacttt gtgcgtactg
301 ctggttacgg gcaagggtca cgcgtggata cggagcatta tctga

```

This corresponds to the amino acid sequence <SEQ ID 812; ORF 227.ng>:

```

g227.pep
1  MNIIRALLII LGCLAAGETA VFLAGIKLPG SIVGMGVLF LQAGWLKTS
51  WLQQLTDALM ANLTLFLVPP CVAVISYLDL IADDWFSLV SASASTLCV
101 LVTGKVHRWI RSII*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 813>:

```

m227.seq (partial)
1  ..ACGTCTTtGC TGCAACAGCT TACCGACGCG CTGATGTCGA ACCTGACGCT
51  GTtCCTCGTG CCgCC.TGCG TGGCGGTCAT CAGCTATTtG GATTtGATTG
101 CCGACGATTG GTTTTCGATA CTGGTTTCCG CCTCCGCCAG cACTTTGTGC
151 GTACTGCTGG TTACGGGCAA AGTCCACCGG TGGATACGGG GTATTATCCG
201 ATGA

```

This corresponds to the amino acid sequence <SEQ ID 814; ORF 227>:

```

m227.pep (partial)
1  ..TSXLQQLTDA LMSNLTFLV PPCVAVISYL DLIADDWFSI LVSASASTLC
51  VLLVTGKVHR WIRGIIR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 227 shows 95.5% identity over a 66 aa overlap with a predicted ORF (ORF 227.ng) from *N. gonorrhoeae*:

```

m227/g227

              10      20      30
m227.pep      TSXLQQLTDALMSNLTFLVPPCVAVISYL
a227          TSXLQQLTDALMSNLTFLVPPCVAVISYL
              10      20      30      40      50      60      70
g227          TAVFLAGIKLPGSIVGMGVLFALLQAGWLKTSWLQQLTDALMANLTLFLVPPCVAVISYL
              20      30      40      50      60      70

```

517

```

a228.pep
1  MKKLLIAAMM AAALAACSQE AKQEVKEAVQ AVESDVKDTA ASAAESAASA
51  VEEAKDQVKD AAADAKASAE EAVTEAKEAV TEAAKDTLNLK AADATQEAAD
101 KMKDAAK*

```

m228/a228 100.0% identity in 107 aa overlap

```

m228.pep      10      20      30      40      50      60
MKKLLIAAMMAAALAACSQEAKQEVKEAVQAVESDVKDTAASAAESAASAVEEAKDQVKD
|||||
a228          10      20      30      40      50      60
MKKLLIAAMMAAALAACSQEAKQEVKEAVQAVESDVKDTAASAAESAASAVEEAKDQVKD

m228.pep      70      80      90      100
AAADAKASAE EAVTEAKEAVTEAAKDTLNLK AADATQEAADKMKDAAKX
|||||
a228          70      80      90      100
AAADAKASAE EAVTEAKEAVTEAAKDTLNLK AADATQEAADKMKDAAKX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 821>:

```

g229.seq
1  atggctgccc tatcgggcgg cggtgcggtc ttcctgataa tgcttccaca
51  tattgcccgc gtccagcgtc agccgccagc gtccgcccaa gcgtcgggag
101 aaatcggcac tgaagccgcc ggcgaaattg tatcggtgct cgcccgaagag
151 gttttgcccg acaaacggca cggtgccgaa cgagcgcgtt accgaacggt
201 tttgatggcc gaacgacagg cgcaggttct gtccgctgaa atctttgtta
251 tcccaataat gcacgccgcg gctgatgccg ccgtagagga aatgatgccc
301 gcccgattg atttcgcgcg acacgcccaa gccgtagcgc aaaccggtgtg
351 ccttttgccg caggctgtcg gcggttttcg tccagcttct gcccgcaaat
401 tcaatcgttt ttccggacga agcgttggtt atagcggatt aacaaaaatc
451 aggacaaggc ggcgggcccg aggcagtacg gatggtacgg aaccggttcg
501 ccggtgctt ggacgcctta gggaaaccgt ccctttgagc cggggcgggg
551 caaccgctac cggtttttgt tcatccgcca tattgtgttg a

```

This corresponds to the amino acid sequence &lt;SEQ ID 822; ORF 229.ng&gt;:

```

g229.pep
1  MAAVSGGAV FLIMLPHIAR VQRPPAFAQ ASGEIGIEAA GEIVSAAAOE
51  VLPDKRHGAE RARYRTVLM AERQAQVLFAE IFVIPIMHAA ADAAVEEMMP
101 ARIDFARHAQ AVAQTVCLLR QAVGGFRPAS ARKFNRFFGR SVVYSGLTKI
151 RTRRRAAGST DGTEPVRPVL GRLREPFPLS RGGATRTGFC SSAILC*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 823>:

```

m229.seq (partial)
1  ..GCTCAAGCGT TGGGAGAAAT CGGCATTGAA GCCGCCGACG AAATGTATC
51  GGCTGCCGCC TAAGAGGTTT TGCTCGACAA ACGGCACGAT GCCGAACGAG
101 CGCGTTACCG AACGGTTTTT ATAGCCGAAC GACAGGCGCA GGCTCTGTTT
151 GCTGAAATCT TTGTTATCCC AATAATGCAC GCCGCCGCGC CTGATGCCGC
201 CGTAGAGGAA ATGATGCCTG CCCGCATTGA TTTCGCGCGA CACGCCTAAG
251 CCCTAGCGCA AACCGTGTGC CTTTTCGCGC AGGCTGTCGG CGGTTTTCGT
301 CCAGCTTCTG CCCGCAAATT CAATCGTTTT TTCGGACGAA GCGTTGTTTA
351 TAGCGGATTA ACAAAAATCA GGACAAGGCA ACGAAGCCGC AGACAGTACA
401 AATAGTACGG AACCGATTCA CTTGGTGCTT CAGCACcTTA GAGAATCGTT
451 CTCTTTTTTG TTCATCCGCT ATATTGTGTT GA

```

This corresponds to the amino acid sequence &lt;SEQ ID 824; ORF 229&gt;:

```

m229.pep (partial)
1  ..AQALGEIGIE AADEIVSAAA XEVLLDKRHD AERARYRTVF IAERQAQALF
51  AEIFVIPIMH AAAADA AVEE MPPARIDFAR HAXALAQTVL LLRQAVGGFR
101 PASARKFNRF FGRSVVYSGL TKIRTRQRSA DSTNSTEPIH LVLQHLRESR
151 SLFCSSAILC *

```

Computer analysis of this amino acid sequence gave the following results:

519

```

          100      110      120      130      140      149
m229.pep  RQAVGGFRPASARKFNRRFFGRSVVYSGLTkIRTRQRSADSTNSTEPIHLVLQHLRES---
          |||||  |||||  |||||  |||||  |||||  |||||
a229      RQAVGGFRPASACKFNRRFFGRSVVYSGLTkIRTRRRSADSTDSTEPIHLVLQHLRESSLX
          130      140      150      160      170      180

          150      160
m229.pep  -----RSLFCSSAILCX
          |: ||||:|
a229      AKARQRRTGFCSSSTI
          190

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 827>:

```

g230.seq
1  atgttcatt ccacgaaaa atacagaaca cccgcccaag tcttattagg
51  cctgattgca ttaacttttg tcggcttcgg cgtcagcacg gtttcccatc
101 cgggcgcgca ctacatcgtc caagtgggcg acgaaaaaat cagcgagcac
151 tcaaaccaaca acgccatgca gaacgagcag gcggacggcg gcagcccttg
201 gcgcgacgcg gtgttccaat ccctgctgca acgcgcctac ctgaaacagg
251 gcgcgaaagt gatgggcatt tcggtttctt ccgaacaaat caagcagatg
301 attgtggacg atcccaattt ccacgacgca aacggcaaat tcagtcacgc
351 gcttttgagt caatacctgt cgcaacgcca tatgtctgaa gaccagtttg
401 tcgaagaaat ccgcgatcag tttgccttgc agaatttggt aagcctcgtc
451 caaacggcg tattggtcgg cgacgcgcag gcggaacagc tgatcagggt
501 gacgcaggtc aaccgcacca tccgttcgca cactttcaac cccgacgagt
551 tcatcgccca agtcaaagcg tctgaagccg atttgcagaa attttataat
601 gcgaacaaaa aagactatct gctgcccagc gcggtcaaat tggaaatatgt
651 cgctttgaat ctgaaggatt ttgcagacaa gcagaccgtc agtgaacagg
701 aagtgaaaaa tgcgtttgaa gagcgcggtg gcggtttgcc ggcacatgaa
751 gccaaacctt ctttcgagca ggaaaaagcc gccgtcgaaa acgaattgaa
801 aatgaaaaaa gcggttgccg acttcaacaa ggcaaaaaga aagctgggcg
851 acgatgcgtt caatcatccc tcctcgcttg ccgaagccgc caaaaacagc
901 ggtttgaaag tggaaaccca agaaacttgg ctgagcaggc aggcgcaca
951 aatgtccggc atgcccgaaa acctaataaa tgccgtattc agcgacgacg
1001 tattgaagaa aaaacacaat tccgaagtgc tgaccatcaa cagcgaaacc
1051 gcgtgggtcg tccgcgccaa agaagtccgc gaagaaaaaa acctactgtt
1101 tgaagaagcc aaagatgcgg tgcgtcaggc ctatatccgt accgaagccg
1151 caaaactttt gaaaacaatg taa

```

This corresponds to the amino acid sequence <SEQ ID 828; ORF 230.ng>:

```

g230.pep
1  MFHSIEKYRT PAQVLLGLIA LTFVGFVST VSHPGADYIV QVGDEKISEH
51  SINNAMQNEQ ADGGSPWRDA VFQSLQRAY LKQGAKLMI SVSSEQIKQM
101 IVDDPNFHDA NGKFSHALLS QYLSQRHMSE DQFVEEIRDQ FALQNLVSLV
151 QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN PDEFIAQVKA SEADLQKFYN
201 ANKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPAHE
251 AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301 GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVL TINSET
351 AWWVRKEVR EEKNLLFEEA KDAVRQAYIR TEAAKLLKTM *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 829>:

```

m230.seq (partial)
1  ATGTTCCATT CCATCGAAAA ATACAGAACG CCCGCCCAAG TCCTTTTGGG
51  CCTGATTGCA TTAACCTTCG TCGGCTTCGG GGTGAGCACG GTATCCCATC
101 CGGGTGCCGA CTACATCGTC CAAGTGGGCG ACGAAAAAaT CAGCGACCAC
151 TCCATCAACA ACGCCATACA GAACGAACAG GCGGACGGCG GCGGCCCTTC
201 GCc.GACGCG GTGTTCCAAT CCCTGCTGCA ACGCGCTAC CTGAAACAGG
251 GCGCGAAGCT GATGGGCATT TCGGTTTCTT CCGAACAAAT CAAGCAAATT
301 ATCGTGGACG ATCCCAATTT CCACGACGCA AACGGCAAAT TCGACCACGC
351 GCTTTTAAAC CGCTACCTTT CCAACGCCA TATGTCTGAA GACCAAGTTG
401 TCGAAGAAAT CCGCGATCAG TTTGCCTTGC AGAATTGTTT AAACCTCGTC

```

521

	310	320	330	340	350	360
m230.pep	GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSVLTINSETAWVVRKEVR					
g230	GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSVLTINSETAWVVRKEVR					
	310	320	330	340	350	360
	370	380				
m230.pep	EEKTLPFAEAKDAVRQAYIRTEAAKL					
g230	EEKNLLFEEAKDAVRQAYIRTEAAKLLKTM					
	370	380	390			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 831>:

**a230.seq** (partial)

```

1  ATGTTCCATT  CCATCGAAAA  ATACAGAACG  CCGGCCCAAG  TCCTTTTGGG
51  CCTGATTGCA  TTAACCTTCG  TCGGCTTCGG  GGTCAGCAGC  GTATCCCATC
101  CGGGTGCCGA  CTACATCGTC  CAAGTGGGCG  ACGAAAAAAT  CAGCGACCAC
151  TCCATCAACA  ACGCCATACA  GAACGAACAG  GCGGACGGCG  GCGGCCCTTC
201  GCGCGACGCG  GTGTTCCAAT  CCCTGCTACA  ACGCGCCTAC  CTGAAACAGG
251  GCGCGAAGCT  GATGGGCATT  TCGGTTTCTT  CCGAACAAAT  CAAGCAGATT
301  ATCGTGGACG  ATCCCAATTT  CCACACGCA  AACGGCAAAT  TCGACCACGC
351  GCTTTTAAAC  CGCTACCTTT  CCCAACGTCA  TATGCTGAA  GACCAGTTTG
401  TCGAAGAAAT  CCGCGATCAG  TTTGCCTTGC  TGAATTTGCT  AAACCTCGTC
451  CAAAACGGCG  TATTGGTCGG  CGACGCGCAG  CCGGAACAGC  TGATCGAGCT
501  GACGCAGGTC  AACCACACCA  TCCGTTTCGA  CACTTTCAAC  CCCGACGAAT
551  TCATCGCCCA  AGTCAAAGTG  TCTGAAGCCG  ATTTGCAGAA  GTTTTATAAC
601  GCAAACAAAA  AAGACTACCT  GCTTCCCAAA  GCGGTCAAAT  TGGAATATGT
651  CGCCTTGAAT  CTGAAAGACT  TTGCAGACAA  ACAGACCGTC  AGCGAAACAG
701  AAGTGAAAAA  TGCGTTTGAA  GAGCGCGTGG  CGCGTTTGCC  GGCAAATGAA
751  GCCAAACCTT  CTTTCGAGCA  GGAAAAAGCC  GCCGTCGAAA  ACGAATTGAA
801  AATGAAAAAG  GCGGTTGCCG  ACTTCAATAA  GGCAAAGAA  AAGCTGGGCG
851  ATGACGCGTT  CAACCATCCT  TCCTCGCTTG  CCGAAGCCGC  CAAAAACAGC
901  GGTTTGAAAG  TCGAAACCCA  AGAACTTGG  CTGAGCAGGC  AGGATGCGCA
951  AATGTCGGGT  ATGCCCGAAA  ACCTGATCAA  TGCCGTATTC  AGCGACGACG
1001  TATTGAAGAA  AAAACACAAT  TCCGAAGTGC  TGACCATCAA  CAGCGAAACC
1051  CGCTGGGTCG  TCCGCGCCAA  AGAAGTCCGC  GAAGAGAAAA  CCTGCGCGTT
1101  TGCCGAAGCC  AAAGACGCGG  TACGTCAGGC  TTATATCCGT  ACCGAAGCCG
1151  CCAAACCTT

```

This corresponds to the amino acid sequence <SEQ ID 832; ORF 230.a>:

**a230.pep** (partial)

```

1  MFHSIEKYRT PAQVLLGLIA LTFVGFVSTV VSHPGADYIV QVGDEKISDH
51  SINNAIQNEQ ADGGGSPSRDA VFQSLQRAY LKQGAALMGI SVSSEQIKQI
101  IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151  QNGVLVGDAQ AEQLIRLTQV NRTIRSHTFN PDEFIAQVKV SEADLQKFYN
201  ANKKDYLLPK AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251  AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDFAFNHP SSLAEAAKNS
301  GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHNSVLTINSET
351  AWWVVRKEVR EEKTLPFMEA KDAVRQAYIR TEAAKL

```

**m230/a230** 99.2% identity in 386 aa overlap

	10	20	30	40	50	60
m230.pep	MFHSIEKYRTPAQVLLGLIALTFVGFVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ					
a230	MFHSIEKYRTPAQVLLGLIALTFVGFVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ					
	10	20	30	40	50	60
	70	80	90	100	110	120
m230.pep	ADGGGSPDAVFQSLQRAYLKQGAALMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN					
a230	ADGGGSPDAVFQSLQRAYLKQGAALMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN					
	70	80	90	100	110	120
	130	140	150	160	170	180

523

```

151 QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN PDEFIAQVKA SEADLQKFYN
201 ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPAHE
251 AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301 GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351 AWVVRAKEVR EEKNLLFEEA KDAVRQAYIR TEAAKLAENK AKEVLTQLNG
401 GKAVDVKWSE VSVLGAQQAR QSMPEAYAE LLKAKPANGK PAYVRLTGLP
451 APVIVEAQAV TPPEDIAAQL FPAKQALAQQ QSANTFDLLI RYFNGKIKQT
501 KGAQSVNDNG GQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 835>:

m230-1.seq

```

1 ATGTTCCATT CCATCGAAAA ATACAGAACG CCCGCCCAAG TCCTTTTGGG
51 CCTGATTGCA TTAACCTTCG TCGGCTTCGG GGTACGACG GTATCCCATC
101 CGGGTGCCGA CTACATCGTC CAAGTGGGCG ACGAAAAAT CAGCGACCAC
151 TCCATCAACA ACGCCATACA GAACGAACAG GCGGACGGCG GCGGCCCTTC
201 GCGCGACGCG GTGTTCCAAT CCCTGCTGCA ACGCGCCTAC CTGAAACAGG
251 GCGCGAAGCT GATGGGCATT TCGGTTTCTT CCGAACAAAT CAAGCAAATT
301 ATCGTGGACG ATCCCAATTT CCAGCAGCGA AACGGCAAAAT TCGACCACGC
351 GCTTTTAAAC CGCTACCTTT CCAACGCCA TATGTCTGAA GACCAGTTTG
401 TCGAAGAAAT CCGCGATCAG TTTGCCTTGC AGAATTTGGT AAACCTCGTC
451 CAAACGCGCG TATTGGTCGG CGACGCGCAG GCGGAACAGC TGATCAGGCT
501 GACACAGGTC AACCGCACCA TCCGTTTCGA CACTTTCAAC CCCGACGAGT
551 TCATCGCCCA AGTCAAAGTG TCTGAAGCCG ATTTGCAGAA ATTTTATAAT
601 GCGAACAAAA AAGACTATCT GTCGCCGCG GCGGTCAAAT TGGAAATATG
651 CGCCTTGAAT CTGAAGGATT TTGCAGACAA GCAGACCGTC AGTGAACCGG
701 AAGTGAAAAA TGCATTTGAA GAGCGCGTGG CGCGTTTGCC GGCAAATGAA
751 GCCAAACCTT CTTTCGAGCA GGAAGAAAGC GCCGTGCAAA ACGAATTGAA
801 AATGAAAAAG GCGGTTGCCG ACTTCAACAA GGCAAAAGAA AAATTGGGCG
851 ACGATGCGTT CAACCATCCT TCCTCGCTTG CCGAAGCCGC CAAAAACAGC
901 GGTTTGAAAG TCGAAACCCA AGAACTTGG CTGAGTAGGC AGGACGCGCA
951 AATGTCGGGT ATGCCCCGAA ACCTGATCAA TGCCGTATTG AGCGACGACG
1001 TATTGAAGAA AAAACACAAT TCCGAAGTGC TGACCATCAA CAGCGAAACC
1051 GCGTGGGTCG TCCGCGCCAA AGAAGTCCGC GAAGAGAAAA CCTTGCCGTT
1101 TGCCGAAGCC AAAGACGCGG TACGTCAGGC TTATATCCGT ACCGAAGCCG
1151 CCAAACCTTG CGAAAAACAAG GCAAAAGACG TGCTTACCCA ACTGAACGGC
1201 GGCAAGGCTG TTGACGTGAA ATGGTCGGAA GTGTCCGTTT TGGGCGCACA
1251 GCAGGCAAGG CAGTCCATGC CGCCCGAGGC TTATGCGGAA CTGCTGAAAG
1301 CAAACCGGCG AAACGGCAAA CCCGCCTACG TCAGGCTGAT CGGCTGCGCG
1351 GCACCCGTGA TTGTCGAAGT ACAGGCTGTA ACCCGCGCGG ATGATATCGC
1401 CGCACAGCTT CCGCTTGCAA AACAGGCTTT GGCAGCAACG CAGTCTGCCA
1451 ATACTTTCGA CTTGTTGATA CGTTATTTC ACGGCAAAAT CAAACAGACC
1501 AAAGGAGCGC AATCGGTGCA CAACGGCGAC GGTCAAGTAA

```

This corresponds to the amino acid sequence <SEQ ID 836; ORF 230-1>:

m230-1.pep

```

1 MFHSIEKYRT PAQVLLGLIA LTFVGFVST VSHPGADYIV QVGDEKISDH
51 SINNAIQNEQ ADGGGPRSDA VFQSLQRAY LKQGAQLMGI SVSSEQIKQI
101 IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151 QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN PDEFIAQVKV SEADLQKFYN
201 ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251 AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301 GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351 AWVVRAKEVR EEKTLPPAEA KDAVRQAYIR TEAAKLAENK AKDVLTQLNG
401 GKAVDVKWSE VSVLGAQQAR QSMPEAYAE LLKAKPANGK PAYVRLIGLP
451 APVIVEVQAV TPPDDIAAQL FLAKQALAQQ QSANTFDLLI RYFNGKIKQT
501 KGAQSVNDNG GQ*

```

m230-1/g230-1 96.3% identity in 512 aa overlap

```

10 20 30 40 50 60
m230-1.pep MFHSIEKYRTPAQVLLGLIALTFVGFVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ
101 IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151 QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN PDEFIAQVKV SEADLQKFYN
201 ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251 AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301 GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351 AWVVRAKEVR EEKTLPPAEA KDAVRQAYIR TEAAKLAENK AKDVLTQLNG
401 GKAVDVKWSE VSVLGAQQAR QSMPEAYAE LLKAKPANGK PAYVRLIGLP
451 APVIVEVQAV TPPDDIAAQL FLAKQALAQQ QSANTFDLLI RYFNGKIKQT
501 KGAQSVNDNG GQ*
70 80 90 100 110 120
m230-1.pep ADGGGPRSDAVFQSLQRAYLKQGAQLMGISVSSEQIKQIIVDDPNFHDANGKFDHALLN
101 IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151 QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN PDEFIAQVKV SEADLQKFYN
201 ANKKDYLLPQ AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251 AKPSFEQEKA AVENELKMKK AVADFNKAKE KLGDDAFNHP SSLAEAAKNS
301 GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351 AWVVRAKEVR EEKTLPPAEA KDAVRQAYIR TEAAKLAENK AKDVLTQLNG
401 GKAVDVKWSE VSVLGAQQAR QSMPEAYAE LLKAKPANGK PAYVRLIGLP
451 APVIVEVQAV TPPDDIAAQL FLAKQALAQQ QSANTFDLLI RYFNGKIKQT
501 KGAQSVNDNG GQ*
130 140 150 160 170 180
m230-1.pep RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN

```

525

This corresponds to the amino acid sequence &lt;SEQ ID 838; ORF 230-1.a&gt;:

## a230-1.pep

```

1 MFHSIEKYRT PAQVLLGLIA LTFVGFVST VSHPGADYIV QVGDEKISDH
51 SINNAIQNEQ ADGGGPSRDA VFQSLQRAY LKQGA KLMI SVSSEIQKI
101 IVDDPNFHDA NGKFDHALLN RYLSQRHMSE DQFVEEIRDQ FALQNLVNLV
151 QNGVLVGDAQ AEQLIRLTQV NRTIRSHFTN PDEFIAQVKV SEADLQKFYN
201 ANKKDYL LPK AVKLEYVALN LKDFADKQTV SETEVKNAFE ERVARLPANE
251 AKPSFEQEKA AVENELMKKK AVADFNKAKE KLGD DAFNHP SSLAEAAKNS
301 GLKVETQETW LSRQDAQMSG MPENLINAVF SDDVLKKKHN SEVLTINSET
351 AWVVRKEVR EEKTL PFAEA KDAVRQAYIR TEAAKLAENK AKDVL TQLNG
401 GKAVDVKWE VSVLGAQQAR QSMPEAYAE LLKAKPANGK PAYVRLIGLP
451 APVIVEQAV TPPDDIAAQL PLAKQALAAQ QSANTFDLLI RYFNGKIKQT
501 KGAQSVNDGD GQ*

```

## a230-1/m230-1 99.8% identity in 512 aa overlap

	10	20	30	40	50	60
a230-1.pep	MFHSIEKYRTPAQVLLGLIALTFVGFVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ					
m230-1	MFHSIEKYRTPAQVLLGLIALTFVGFVSTVSHPGADYIVQVGDEKISDHSINNAIQNEQ					
	10	20	30	40	50	60
	70	80	90	100	110	120
a230-1.pep	ADGGGPSRDAVFQSLQRAYLKQGA KLMI SVSSEIQKIIVDDPNFHDA NGKFDHALLN					
m230-1	ADGGGPSRDAVFQSLQRAYLKQGA KLMI SVSSEIQKIIVDDPNFHDA NGKFDHALLN					
	70	80	90	100	110	120
	130	140	150	160	170	180
a230-1.pep	RYLSQRHMSEDQFVEEIRDQFALQNLVNLVQNGVLVGDAQAEQLIRLTQVNRIRSHFTN					
m230-1	RYLSQRHMSEDQFVEEIRDQFALQNLVNLVQNGVLVGDAQAEQLIRLTQVNRIRSHFTN					
	130	140	150	160	170	180
	190	200	210	220	230	240
a230-1.pep	PDEFIAQVKVSEADLQKFYNANKKDYL LPKAVKLEYVALNLKDFADKQTVSETEVKNAFE					
m230-1	PDEFIAQVKVSEADLQKFYNANKKDYL LPKAVKLEYVALNLKDFADKQTVSETEVKNAFE					
	190	200	210	220	230	240
	250	260	270	280	290	300
a230-1.pep	ERVARLPANEAKPSFEQEKA AVENELMKKKAVADFNKAKEKLGD DAFNHPSSLAEAAKNS					
m230-1	ERVARLPANEAKPSFEQEKA AVENELMKKKAVADFNKAKEKLGD DAFNHPSSLAEAAKNS					
	250	260	270	280	290	300
	310	320	330	340	350	360
a230-1.pep	GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSEVLTINSETAWVVRKEVR					
m230-1	GLKVETQETWLSRQDAQMSGMPENLINAVFSDDVLKKKHNSEVLTINSETAWVVRKEVR					
	310	320	330	340	350	360
	370	380	390	400	410	420
a230-1.pep	EEKTL PFAEAKDAVRQAYIRTEAAKLAENKAKDVL TQLNGGKAVDVKWEVSVLGAQQAR					
m230-1	EEKTL PFAEAKDAVRQAYIRTEAAKLAENKAKDVL TQLNGGKAVDVKWEVSVLGAQQAR					
	370	380	390	400	410	420
	430	440	450	460	470	480
a230-1.pep	QSMPEAYAE LLKAKPANGKPAYVRLIGLPAPVIVEQAVTPDDIAAQLPLAKQALAAQ					
m230-1	QSMPEAYAE LLKAKPANGKPAYVRLIGLPAPVIVEQAVTPDDIAAQLPLAKQALAAQ					
	430	440	450	460	470	480
	490	500	510			
a230-1.pep	QSANTFDLLIRYFNGKIKQTKGAQSVNDGDGQX					
m230-1	QSANTFDLLIRYFNGKIKQTKGAQSVNDGDGQX					
	490	500	510			



527

**a231.seq** (partial)

```

1 ATGTCAAAAC GAAAATCCAT AAACCGTCCG TATCAAAAAC CGGCGGAACT
51 GCCGCCGTG CAAATAATC CGCCATTTTA CCGTAAAAAC CGCCGCTGA
101 ACTTTTTTAT CGNGGCAGAC GGCGGTTGCG CGTCTCCGCA AAAATGCAGG
151 GCGCGCGGTT TTCAGACGGC ATTTGCCGTT CAAAGCCGTG CGGTGTCTTT
201 ACCAAATGCC CAACCATTCG CCCACGGCAT CCATCCAATC CTTATTGCCC
251 CCGCCGCTCC TGCCTGCCCG GCGGTACGCC CACGGCGCTT GCGGATTTT
301 AGCTTCCAC AATCCTTTGC GTTCCCTTC CGCCTGAATT TGAGCGTCGG
351 CATAATCGGC AAAATCCGCC TTATCCTGCT GTTCTTTAGC ATAACITTTA
401 TAATGCCACG CCGCCCCGTC CTGCACCTGC ATCAGGTTCA AATCGGTTTT
451 GCCGACAGAA ACCTGCGCCA CTTCGCGCTG GTAGCGGTCG GTGTCAACA
501 CGCGGACGCT GACTTTCCTG CCTTCGCGC CGCGCGCAG GTTGTCCGC
551 GAACGCGTGC CGTAAGCCTG TTTCATCTCC GGCGCGTCGA TATACGCCAT
601 CCGGATTTTG TGTTTCGCGC CGTCGCGCTC GATAACGTGA AGGGTGTGCG
651 CGTCATAGAC TTTGGACACC GTGCCTGTGT AGCGGTGGCC GGATTTCGCC
701 GATGCTCGGC GGCGGGCGGG CGGTCGGAA CCGCGCTCCC CTGCGCGCC
751 GAGTACGTCG AGTACGGCAA CCGCCGTCG CACCGCCTCG CTGCCGTACC
801 CCGTATAACC CAACGCACCC AAAAGCGACA AGGCGACGGG AAGCCATTT
851 ATGATTTTTT TAATCTGCAT ATTTTCAA TGCCGATGCC GTCTGAACAT
901 ATC

```

This corresponds to the amino acid sequence <SEQ ID 844; ORF 217.a>:

**a231.pep** (partial)

```

1 MSKRKSINRP YQKPAELPPL QNNPPFYRKN RRLNFFIXAD GGCASPQKCR
51 ARGFQTAFV QSRVSLPNA QPFAHGIHPI LIAPAAPACP AVRPRRLRIF
101 SFPQSFAFFP RLNLVSGIIG KIRLILFFS ITFIMPRRPV LHLHQVQIGF
151 ADRNLRHFAL VAVGVEHADA DFFAFRRRAQ VVARTRAVSL FHLRRVDIRH
201 PDFVFRVAV DNVKGVAVID FGHRACVAVA GFRRCSAAGG RVGTRVPCRA
251 EYVEYGNRRP HRLAAVPRIT QRTQKRQGDG KPFHDFENLH IFQMPMPSEH
301 I

```

**m231/a231** 98.6% identity in 73 aa overlap

	10	20	30	40	50	60
m231.pep	MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIAADGGCASPQKCRARGFQTAFV					
a231	MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIXADGGCASPQKCRARGFQTAFV					
	10	20	30	40	50	60
	70					
m231.pep	QSRVSLPNAQPFQ					
a231	QSRVSLPNAQPFQFAHGIHPILIAPAAPACPAVRPRRLRIFSFPQSFAFFPRLNLVSGIIG					
	70	80	90	100	110	120

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 845>:

**g231-1.seq**

```

1 ATGTCAAAAC GAAAATCCAT AAACCGTCCG TATCAAAAAC CGGCGGAACT
51 GCCGCCGTG CAAAATAATC CGCCATTTTA CCGTAAAAAC CGCCGCTGA
101 ACTTTTTTAT CGCGGCAGAC GGCGGTTGCG CGTCTCCGCA AAAATGCAGG
151 GCGCGCGGTT TTCAGACGGC ATTTGCCGTT CAAAGCCGTG CGGTGTCTTT
201 ACCAAATGCC CAACCATTCG CCCACGGAAT CCATCCAATC CTTATTGCCC
251 CCGCCGCTCC TGCCTGCCCG GCGGTACGCC CACGGCGCTT GCGGATTTT
301 AGCTTCCAC AATCCTTTGC GTTCCCTTC CGCCTGAATT TGAGCGTCGG
351 CATAGTCGGC AAAATCCGCC TTATCCTGCT GTTCTTTAGC ATAACITTTA
401 TAATGCCACG CCGCCCCGTC CTGCACCTGC ATCAGGTTCA AATCGGTTTT
451 GCCGCGGAT ACCTGCGCCA CTTCGCGCTG ATAGCGGTCG GTTTCAAACa
501 CaCgTaCaat gagtttcgtA ccctccGCCG ccgcgcgCAG GTTgtcgcGC
551 GAACgTGtAC CGTAagcgtg Ttcatctcc GGTGcgtcGA TATACGCCaT
601 cCgAATTTta tGtTtcgcgc cgtcgcCgtc gATGACGTGA AGGgtatcGC
651 CgtcATAGAC TTTGGACACC Gtgcctgcgt AGctGTGGCC GGATtctgc

```

This corresponds to the amino acid sequence <SEQ ID 846; ORF 231-1.ng>:

**g231-1.pep**

```

1 MSKRKSINRP YQKPAELPPL QNNPPFYRKN RRLNFFIAAD GGCASPQKCR
51 ARGFQTAFV QSRVSLPNA QPFAHGIHPI LIAPAAPACP AVRPRRLRIF

```

529

```

151 GCGCGCGGTT TTCAGACGGC ATTTGCCGTT CAAAGCCGTG CCGTGTCTTT
201 ACCAAATGCC CAACCATTCG CCCACGGCAT CCATCCAATC CTTATTGCCC
251 CCGCCGCTCC TGCTTCCCG GCGGTACGCC CACGGCGCTT GCGGATTTT
301 AGCTTTCAC AATCCTTTC GTTCCCTTTC CGCTGAATT TGAGCGTCGG
351 CATAATCGGC AAAATCCGCC TTATCCTGCT GTTCTTTAGC ATAACTTTTA
401 TAATGCCACG CCGCCCCGTC CTGCACCTGC ATCAGGTTC AATCGGTTT
451 GCGGACAGAA ACCTGCGCCA CTTCCGCTG GTAGCGGTC GTGTCGAACA
501 CCGGAGCGCT GACTTTCCTG CCTTCCGCG CCGCGCGCAG GTTGTGCGCG
551 GAACGCGTGC CGTAAGCCTG TTTTCATCTC GCGCGCTCGA TATACGCCAT
601 CCGGATTTTG TGTTTCGCGC CGTCCGCTG GATAACGTGA AGGTGTGCGC
651 CGTCATAGAC TTTGGACACC GTGCCTGTGT AGCGGTGGCC GGATTTCGCC
701 GATGCTCGGC GCGGGCGGG CGCGTCGGA CCCGCTGCC CTGCCGCGCC
751 GAGTACGTCG AGTACGGCAA CCGCGTCCG CACCGCTCG CTGCCGTACC
801 CCGTATAACC CAACGCACCC AAAAGCGACA AGCGACGGG AAGCCATTT
851 ATGATTTTTT TAATCTGCAT ATTTTCAAA TGCCGATGCC GTCTGAACAT
901 ATCGGAATCG GATTTCAGAC GGCATCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 850; ORF 231-1.a>:

**a231-1.pep**

```

1 MSKRKSINRP YQKPAELPPL QNNPPFYRKN RRLNFFIAAD GGCASPQKCR
51 ARGFQTAFV QSRVSLPNA QPFAHGIHPI LIAPAACP AVRPRLRIF
101 SFPQSFAFP RLNLVGIIG KIRLILFFS ITFIMPRPV LHLHQVQIGF
151 ADRNLRFAL VAVGVEHADA DPAFRRAQ VVARTRAV FHLRRVDIRH
201 PDFVFRVAV DNVKGVAID FGHRACVAV GFRRCSAAG RVGTRVPCRA
251 EYVEYGNRRP HRLAAVPRIT QRTQKROGDG KPFHDFNLH IFQMPMPSEH
301 IGIGFQTAS*

```

**a231-1/m231-1** 99.0% identity in 309 aa overlap

	10	20	30	40	50	60
a231-1.pep	MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIAADGGCASPQKCRARGFQTAFV					
m231-1	MSKRKSINRPYQKPAELPPLQNNPPFYRKNRRLNFFIAADGGCASPQKCRARGFQTAFV					
	10	20	30	40	50	60
	70	80	90	100	110	120
a231-1.pep	QSRVSLPNAQPFAHGIHPI LIAPAACP AVRPRLRIF SFPQSFAFP RLNLVGIIG					
m231-1	QSRVSLPNAQPFAHGIHPI LIAPAACP AVRPRLRIF SFPQSFAFP RLNLVGIIG					
	70	80	90	100	110	120
	130	140	150	160	170	180
a231-1.pep	KIRLILFFSITFIMPRPV LHLHQVQIGFADRNLRFALVAVGVEHADADPAFRRAQ					
m231-1	KIRLILFFSITFIMPRPV LHLHQVQIGFADRNLRFALVAVGVEHADADPAFRRAQ					
	130	140	150	160	170	180
	190	200	210	220	230	240
a231-1.pep	VVARTRAVSLFHLRRVDIRHPDFVFRVAVD DNVKGVAIDFGHRACVAVAGFRRCSAAGG					
m231-1	VVARTRAVSLFHLRRVDIRHPDFVFRVAVD DNVKGVAIDFGHRACVAVAGFRRCSAAGG					
	190	200	210	220	230	240
	250	260	270	280	290	300
a231-1.pep	RVGTRVPCRAEYVEYGNRRPHRLAAVPRITQRTQKROGDGKPFHDFNLHIFQMPMPSEH					
m231-1	RVGTRVPCRAEYVEYGNRRPHRLAAVPRITQRTQKROGDGKPFHDFNLHIFQMPMPSEH					
	250	260	270	280	290	300
	310					
a231-1.pep	IGIGFQTASX					
m231-1	IGIGFQTASX					
	310					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 851>:

**g232.seq**

```

1 atgatgggca acagcctgat tgaatccggt acgtttgtcg ccatcctggt
51 tggtcagatt ttgggaacgg cggttgccgg cgcgcgcct tatattgtcg
101 ggatactggt tttgctgggc gccgtcggag gaacggcgg cagcctgttt

```

531

	10	20	30	40	50	60
	70	80	90	100	110	120
m232.pep	TQIEWNIVRGTKSLLRETVRHKPVFTAIIGISWFWFVGAVYTTQLPTFTQIHLGGNDNVF					
g232	TQIEWNIVRGTKSLLRETVRHNPVFTAIIGISWFWFVGAVYTTQLPTFTQIHLGGNDNVF					
	130	140	150	160	170	180
m232.pep	NLMLALFSIGIAAGSVLCAKFSXERLMLAWVTVGALGLTVCGLVLVWLTHGHRFEGLNGI					
g232	NLMLALFSIGIAAGSVLCAKFGRERLMLAWVTVGALGSTVCGLVLVWLTHGHRFEGLNGI					
	190	200	210	220	230	
m232.pep	FXFLSQGWAYPVMAMVTIGFFGGFFSVPLYT-VQTAIARFPRP-AVAANNIVNGIFMVS					
g232	FWFLSQGWAYPVMAMVTIGFFGGFFSVPLYTWLQTASSETFRARAANAANNIVNGIFMVS					
	240	250	260	270	280	289
m232.pep	AAVLSAVLLFLFDSISLLYLIALGNIPLSVFLIKRERFLGAAAIRKKPX					
g232	AAVLSAVLLFLFDSISLLYLIALGNIPLAFLIKRERFLGAAAIRKKP					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 855>:

```

a232.seq
1  ATGTACGCTA AAAAAGGCGG TTTGGGACTG GTTAAAAGCC GCCGTTTCGC
51  ACCTCTTTTC GCTACGCAGT TTCTCGGCGC GTTCAACGAC AATGTGTTCA
101 AAACCGCGCT GTTTGTGATG ATTGGGTTTT ACGGTTTGGG GCAAAACGGC
151 TTCCTGCCTG CCGGACAGAT GTTGAAC TTG GCGCGTTGC TGTATTATTT
201 GCCGTATTTT CTGCTTTCCT CCGCTGCGGG GCAGTTGGGT AACAAATTCG
251 ACAAGCCGCT TTTGGCGCGT TGGGCCAAGG TGCTGGAAT GATCATTATG
301 GCGGTGGCGG CATACGGGTT TTATATCCGG TCTGCCCCGC TGCTTTTGCG
351 GTGTCTGTTT TGCATGGGCG CGCAATCGAC GCTGTTCCGG CCGCTGAAAT
401 ACGCCATCCT GCCCGATTAT CTCGACGACA AAGAGTTGAT GATGGGCAAC
451 AGCCTGATTG AATCGGGTAC GTTTGTGCGC ATCTGTTCG GTCAGATACT
501 GGGGACTGCG GTGGCAGGTG TACCGCCTTA TATTGTCGGG ATACTGGTTT
551 TGCTGGTCGC CGTAGGAGGC ACGGTCGGCA GCCTGTTTAT GCCGTCCGTA
601 CCCGCCAAGG CTGCCGATAC ACAAAATTGAG TGGAATATTG TCCGGGGTAC
651 AAAATCCCTG CTGCGTGAAG CCGTGCGGCA CAAGCCCGTT TTTACCGCCA
701 TTATCGGTAT TTCGTGTTT TGGTTTGTCG GCGCGGTTTA TACCACGCAA
751 CTGCCGACCT TTACCCAAAT CCATCTAGGC GGCAACGACA ATGTTTTCAA
801 CCTGATGCTT GCCCTGTTT CCATCGGTAT TGCCGCCGGT TCGGTACTGT
851 GTGCCAAGTT CAGCAGGGAA CGGCTGAGGT TGGCTTGGGT AACGGTTGGT
901 GCGTTGGGTT TGACGGTTTG CGGCTTGGTT TTGGTGTGGC TGACGCACGG
951 ACACCGTTTT GAAGGGCTGA ACGGCATTTT TTGGTTTTTA TCGCAAGGAT
1001 GGGCATATCC CGTGATGGCG GTGATGACGC TGATCGGCTT TTTCGGCGGA
1051 TTTTCTCCG TTCCGCTCTA TACCTGGCTG CAAACCGCCA GTAGCGAGAC
1101 TTCCCGCGCC CGCGCGTTG CCGCAACAA TATCGTTAAC GGTATTTTA
1151 TGGTTTCCGC TGCCGTTTGG AGCGCGGTGT TGCTGTTTTT GTTTGACAGC
1201 ATTTCTTGT TGTATCTGAT TGTGCTTTG GGCAATATTC CGTTGTCGST
1251 ATTTTGTATT AAGCGCGAAA GCGGTTTTT AGGCGCGGCG GCAATCAGGA
1301 AAAAACCTTG A

```

This corresponds to the amino acid sequence <SEQ ID 856; ORF 232.a>:

```

a232.pep
1  MYAKKGGLGL VKSRRFAPLF ATQFLGAFND NVFKTALFVM IGFYGLGQNG
51  FLPAGQMLNL GALLFILPYF LFSSLSGQLG NKFDKAVLAR WAKVLEMIIM
101 AVAAYGFYIR SAPLLACL F CMGAQSTLFG PLKYAILPDY LDDKELMMGN
151 SLIESGTFVA ILFGQILGTA VAGVPPYIVG ILVLLVAVGG TVGSLFMPSV
201 PAKAADTQIE WNIVRGTKSL LRETVRHKPV FTAIIGISWF WFGVAVYTTQ
251 LPTFTQIHLG GNDNVFNML ALFSIGIAAG SVLCAKFSRE RLRLAWVTVG

```

151 QISATVDRSG LWQAQTPQLF QAGLLHRLA AENLGGITDE ASAVEKLGV  
 201 PLLIQGDARN LKLTQPDAY IVRLLLNAV\*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 859>:

m233.seq (partial)  
 1 ATGAAGCGCA AAAATATCGC GCTGATTCCC GCCGCCGGCA TCGGGGCGCG  
 51 TTTCGGTGCG GACAAACCCA AGCAATATGT CGAAATCGGA AGCAAAACCG  
 101 TTTTAGAACA TACGATTGGG ATTTTGAAC GGCATGAGGC CGTCGATTG  
 151 ACCGTCGTTG TCGTCTCGCC CGAAGACACG TTTGCCGATA AGGTTTCAGAC  
 201 GGCATTTCCT CAGGTTTCGGG TGTGGAAAAA CGGCGGACAG ACCCGCGCCG  
 251 AAACCGTCCG CAACGGTGTG GCAAACTGT TGGAAACCGG TTTGGCGGCG  
 301 GAAACCGACA ATATTCTGGT ACACGATGCC GCGCGTTGCT GCCTGCCGTC  
 351 TGAAGCTTTG ACGCGGTTGA TAGAACAGGC GGGCAACGCC GCCGAAGGCG  
 401 GGATTTTGGC AATCCCATT GCCGATACGC TCAAGTGC GC GACGGTGGG  
 451 AACATT...

This corresponds to the amino acid sequence <SEQ ID 860; ORF 233>:

m233.pep (partial)  
 1 MKRKNIALIP AAGIGARFGA DKPKQYVEIG SKTVLEHTIG IFERHEAVDL  
 51 TVVVVSPEDT FADKVQTAFF QVRVWKNNGQ TRAETVRNGV AKLLETGLAA  
 101 ETDNILVHDA ARCLLPSEAL TRLIEQAGNA AEGGILAIPI ADTLKCADGG  
 151 NI...

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 233 shows 93.4% identity over a 152 aa overlap with a predicted ORF (ORF 233.ng) from *N. gonorrhoeae*:

m233/g233

m233.pep	10	20	30	40	50	60
	MKRKNIALIP AAGIGARFGADKPKQYVEIGSKTVLEHTIGIFERHEAVDLTVVVVSPEDT					
g233	MKRKNIALIP AAGIGVRFGADKPKQYVEIGSKTVLEHVLGIFERHEAVDLTVVVVSPEDT					
	10	20	30	40	50	60
m233.pep	70	80	90	100	110	120
	FADKVQTAFFQVRVWKNNGQTRAETVRNGVAKLLETGLAAETDNILVHDAARCLLPSEAL					
g233	FADKVQTAFFQVRVWKNNGQTRAETVRNGVAKLLETGLAAETDNILVHDAARCLLPSEAL					
	70	80	90	100	110	120
m233.pep	130	140	150			
	TRLIEQAGNAAEGGILAIPIADTLKCADGGNI					
g233	ARLIEQAGNAAEGGILAVPVADTLKRAESGQISATVDRSGLWQAQTPQLFQAGLLHRLA					
	130	140	150	160	170	180

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 861>:

a233.seq  
 1 ATGAAGCGCA AAAATATCGC GCTGATTCCC GCCGCCGGCA TCGGGGCGCG  
 51 TTTCGGTGCG GACAAACCCA AGCAATATGT CGAAATCGGA AGCAAAACCG  
 101 TTTTAGAACA TACGATTGGG ATTTTGAAC GGCATGAGGC CGTCGATTG  
 151 ACCGTCGTTG TCGTCTCGCC CGAAGACACG TTTGCCGATA AGGTTTCAGAC  
 201 GGCATTTCCT CAGGTTTCGGG TGTGGAAAAA CGGCGGACAG ACCCGCGCCG  
 251 AAACGTCCG CAACGGTGTG GCAAAATTGT TGGAAACCGG TTTGGCGGCG  
 301 GAAACCGACA ATATTCTGGT ACACGATGCC GCGCGTTGCT GCCTGCCGTC  
 351 TGAAGCTTTG ACGCGGTTGA TAGAACAGGC GGGCAACGCT GCCGAAGGTG  
 401 GGATTTTGGC AATCCCATT GCCGATACGC TCAAGTGC GC GACGGTGGG  
 451 AACATTAGTG CAACCGTCTGA GCGGACGAGC CTTTGGCAGG CGCAACGCC  
 501 GCAGCTTTTC CGCGCCGGGC TGCTGCACCG CGCATTTGGCT CGGAAAACT  
 551 TGGACGGCAT TACCGATGAA GCGTCCGCCG TGGAAAAATT GGGCATCCGC  
 601 CCTTTGCTGG TGCAGGGCGA CGCGCGCAAT TTGAACTGA CGCAGCCGCA  
 651 GGACGCATAC ATCGTCAGGC TGCTGCTCGA TGCCGTCTGA

535

m234.pep (partial)  
 1 ..GAGEYALSNR EIIGFGGTSG YDATLNGKVL DLAIREAVNS LVQAVDNGAW  
 51 QPNR\*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 234 shows 94.4% identity over a 54 aa overlap with a predicted ORF (ORF 234.ng) from *N. gonorrhoeae*:

m234/g234

				10	20	30
m234.pep				GAGEYALSNREI	IGFGGTSGYDATLNGKVL	
g234	LGRGKSQIAYAKVALNIVNVNTSEIVYSTQGAGEYALSNREI	IGFGGTSGYDATLNGKVL				
	140	150	160	170	180	190

		40	50
m234.pep	DLAIREAVNSLVQAVDNGAWQPNRX		
g234	DLAIREAVDNLVQAVDNGAWQSNRX		
	200	210	220

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 867>:

a234.seq (partial)  
 1 AACCGCACCT ATTTGAACGC ATTAACACAG GAATCCGGCA TTTCCGGCAA  
 51 AGCGCATAAC CTGAAAGGCG CAAATTATGT CGNNACCGGC GATGTAACCG  
 101 AATTCGGACG CANAGATGTC GGCGATCATC AGCTCTTCGG CATTTCGGGT  
 151 CGCGGCAAAT CGCAAATCGC CTATGCAAAA GTGGCTCTGA ATATCGTCAA  
 201 CGTCAATACT TCCGAAATCG TCTATTCCGC ACAGGGCGCG GGCGAATACG  
 251 CACTTTCCAA CCGTGAAATC ATCGGTTTCG GCGGCACTTC CGGTACGAT  
 301 GCGACTTTGA ACGGCAAAGT TTTAGACTTG GCAATCCGCG AAGCCGTCAA  
 351 CAGCCTGTTT CAGGCTGTTG ACAACGGCGC ATGGCAACCC AACCGTTAA

This corresponds to the amino acid sequence <SEQ ID 868; ORF 234.a>:

a234.pep (partial)  
 1 NRTYLNALKQ ESGISGAHN LKGANYVXTG DVTEFGRXDV GDHQLFGILG  
 51 RGKSQIAYAK VALNIVNVNT SEIVYSAQGA GEYALSNREI IGFGGTSGYD  
 101 ATLNGKVLDL AIREAVNSLV QAVDNGAWQP NR\*

m234/a234 100.0% identity in 54 aa overlap

				10	20	
30						
m234.pep				GAGEYALSNREI	IGFGGTSGYDATLNGKVL	
a234	LGRGKSQIAYAKVALNIVNVNTSEIVYSAQGA	GEYALSNREI	IGFGGTSGYDATLNGKVL			
	50	60	70	80	90	

		40	50
m234.pep	DLAIREAVNSLVQAVDNGAWQPNRX		
a234	DLAIREAVNSLVQAVDNGAWQPNRX		
	110	120	130

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 869>:

g235.seq

537

```

m235.pep  YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
g235      YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
              130      140      150      160      170      180

              190      200      210
m235.pep  DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
g235      DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
              190      200      210

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 873>:

```

a235.seq
1  ATGAAACCTT TGATTTTAGG GCTTGCCGCC GTGTTGGCGC TGTCTGCCTG
51  CCAAGTCAA  AAAGCGCCCG ATTCGACTA  CACGTCATTC AAGGAAAGCA
101 AACCGGCTTC AATTTTGGTG GTTCCGCCGC TGAACGAATC GCCCGATGTC
151 AACCGAACAT GGGGTGTACT GGCTTCGACC GCCGCGCCGC TTTCGAAGC
201 CGGCTATTAC GTCTTCCCGC CCGCAGTCGT GGAGGAAACC TTCAAACAAA
251 ACGGCTTGAC CAATGCCGCC GATATTCACG CCGTCCGCCG GGAAAACTG
301 CATCAGATT  TCGGCAATGA TCGGTTTGT  TACATTACGG TTACCGAATA
351 CCGCACTTCA TATCAAATT  TAGACAGCGT GACGACCGTA TCCGCCAAAG
401 CACGGCTGGT CGATTCCCGC AACGGAAGAG AGTTGTGGTC GGGTTCGGCC
451 AGCATCCGCG AAGGCAGCAA CAACAGCAAC AGCGGCCTGT TGGGGGCTTT
501 GGTCAGCGCA GTGGTCAATC AGATTGCCA  CAGCCTGACC GACCGCGGTT
551 ATCAGGTTTC TAAACCGGCC GCATACAACC TGCTGTCGCC CTATTCTCAC
601 AACGCATCT  TGAAAGGTCC GAGATTCGTC GAAGAGCAGC CCAATAAA

```

This corresponds to the amino acid sequence <SEQ ID 874; ORF 235.a>:

```

a235.pep
1  MKPLILGLAA VLALSACQVQ KAPDFDYTSF KESKPASILV VPPLNESPDV
51  NGTWGVLAST AAPLSEAGYY VFPAAVVEET FKQNGLTNAA DIHAVRPEKL
101 HOIFGNDAVL YITVTEYGTS YQILDSVTTV SAKARLVDSR NGKELWSGSA
151 SIREGSNNSN SGLLGALVSA VVNQIANSLT DRGYQVSKTA AYNLLSPYSH
201 NGILKGPRFV EEQPK*

```

m235/a235 100.0% identity in 215 aa overlap

```

              10      20      30      40      50      60
m235.pep  MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
a235      MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
              10      20      30      40      50      60

              70      80      90      100     110     120
m235.pep  AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIAHVRPEKLHQIFGNDAVLVITVTEYGTS
a235      AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIAHVRPEKLHQIFGNDAVLVITVTEYGTS
              70      80      90      100     110     120

              130     140     150     160     170     180
m235.pep  YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
a235      YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
              130     140     150     160     170     180

              190     200     210
m235.pep  DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
a235      DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPKX
              190     200     210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 875>:

g236.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 879>:

This corresponds to the amino acid sequence <SEQ ID 880; ORF 236.a>:

```
a236.pep
1 MARFAFSADI LCTAFADGFM ACNRAHIAGV VPAAFAFFTI TGFSNGKFA
51 AYFHFCFRHQ QSKAQFFAQ IGIAGHFFRR GNFGFLQGR TDCGFVGAQRL
101 DGGGYRFAGF ADCRPFFHQF QFGFVFDGRE LVPMSMKHAV FCAAADDVPR
151 FFAGEAQNRC NQENQAARDV VQGGRLAAAG AAVGFGGIEA VFQDIEVERA
201 QVFRAERNHF FHKKVEGIR IKITAGNAFLQ PPCQHQGIIV DFHHIRLLHG
251 IFNRIEVAQV GKQKAQGIAD TAVAFGYALE DFFGNRQFAA VIGGCRPQQAQ
301 DVRAELVIHF LRRDDVADGF RHFAFVLIHH ETMCOOLEVR RATH*
```

541

```

101 AFQRPEYRVQ TDISIDGIDH IFTLDAAFGR VNQPTVLIRF DARHRGLEN
151 LRAGFGSGAG NTQRVIERMK MPQGIELCA LVHIAVKLLF QLGFIPKLM
201 TRTVMLGVF MPLLQLFPML RTDGNRGITA LPIAIDGMFA DAFVHQFDRL
251 QRLLPKPLRL LQADLFFNFP HTAGVVADNL PATPAGRAEA DTCGFQNDGF
301 MSVFRQRQCG TQTAQTAADN AGFGFQTALE FRINSMRINR TKIIRRQIFL
351 KIRANHCVCF IGYIFGRNDT DCRAISSXQK IG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 883>:

```

m237.seq
1 ATGCGGGACA AGGTTGGCGG TAATGTCGCA CTCCCCGCCC CACGAATATT
51 CGATTTTGAC ATCGGCAAGC TGCGGAAAAA CTTTAAGCAT ATCTTGGCGG
101 ACAAGCTCGG TCATaCGCTC AGGATTGTCG ATAAACTCGT TATCCTTACC
151 GCCGAAAAGC AGTCTGCCGT CCGCGCTGAG GCGTAATAA TCCAAAATAT
201 GGCGGTGTGC GCATACTGCC ATATTGTTA GGATAAGCCC TTTTGGCGCG
251 GCCCCAAGG GTCGCTCGC AATAATAAAG GTGCTGACAG CAATCGCCTT
301 GCGTTCCAAA GGCCGGAATA TCGGGTTCAA ACCTGCATAA GTATTGACAG
351 CATAGACCAC ATTTTTCAC TCGACGCTGC CTTGGGGCGT GTAAACCAGC
401 CAACCGTTTT GATGCGGTTT GATGCACGTC ATCGGGGATT GCTCGAAAAT
451 CTGCGCACCG GCTTCGGCAG CGGCACGAGC GATGCCCAA GTGTAAGTGA
501 GCGGATGCAG GTGTCCGGAT AAGGGGTGCA ATTGTGCCCC TTGGTACATA
551 TCGCTGTCAA GCTGCTGTTT CAACTCGGCT TTATCCCAA GTTGATAATG
601 ACTCGCACCG TAATGCCGTT GGGCGTGTTT ATGCCACTGC TGCAACTCTT
651 CCCAATGCTG CGGACGGACG GCAACCGTGG CATAACCGCG CTGCCAATCA
701 CAATCGACCG CATGTTTTCG GACGCGTTCG TCCACCAGT CGACCGCCTG
751 CAAAGACTGT TGCCAAAACC ATTGCGCCTG CTCCAAGCCG ACCTGTTTTT
801 CAATTTCCCC CATACCGCAG nCGTAATCGC TGATAACCTG CCCGCCACTC
851 CGTCCCGACG CGCCGAAAACC GATACGCGCG GCTTCCAACA CAACCGTTTC
901 ATGTCCCTGC TCCGCCAAGG GCAATGCAGT GCACAAACCA CCCAATCCGC
951 CGCCGATGAT ACAGGTATCG GTTTTCAGAC GGCATTGAAG TtCGGATAA
1001 ACAGTATGAG GATTAAACCGA ACTGAAATAA TAAGAAGGCA GATATTCTTG
1051 AAAATCAGGG CGAATCATTG TGTTTGCTTT ATCAGGTGTA TTTTCGGACG
1101 GAATGATACA GGCTGTCGGG CCATATCGTC CAwACAGAAA ATCGGTTGA

```

This corresponds to the amino acid sequence <SEQ ID 884; ORF 237>:

```

m237.pep
1 MRDKVGGNVA LPAPRIFDFD IGKLRKNFKH ILADKLGHTL RIVDKLVILT
51 AEKQSAVRAE AVIIQNMAVV AYCHIVTDKP FCARPQGFGR NNGGADSNRL
101 AFQRPEYRVQ TCISIDSIDH IFALDAAFR VNQPTVLMRF DARHRGLEN
151 LRTGFGSGTS DAQSVSERMQ VSGXGVELCP LVHIAVKLLF QLGFIPKLM
201 TRTVMLGVF MPLLQLFPML RTDGNRGITA LPITIDGMFA DAFVHQFDRL
251 QRLLPKPLRL LQADLFFNFP HTAXVIADNL PATPSRAET DTRGFQHNRF
301 MSLRQGCQS AQTTQSAADD TGIGFQALK FRINSMRINR TEIIRRQIFL
351 KIRANHCVCF IRCIFGRNDT GCRAISSXQK IG*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 237 shows 86.1% identity over a 382 aa overlap with a predicted ORF (ORF 237.ng) from *N. gonorrhoeae*:

```

m237/g237
10 20 30 40 50 60
m237.pep MRDKVGGNVALPAPRIFDFDIGKLRKNFKHILADKLGHTRIVDKLVILTAEKQSAVRAE
|||||:|||||:|||||:|||||:|||||:|||||
g237 MRDKVGGNIALPAPRIFDSNIGKLRKNFKHILADKLGHTRIVDKFVILTAEKQPAVRAE
10 20 30 40 50 60

70 80 90 100 110 120
m237.pep AVIIQNMAVVAYCHIVTDKPF CARPQGFGRNNGGADSNRLAFQRPEYRVQTCISIDSIDH
|||||:|||||:|||||:|||||:|||||:|||||
g237 AVIIQNMAVVAYCHIVADKPF CARAQFGGNNKGADGNRLAFQRPEYRVQTDISIDGIDH
70 80 90 100 110 120

130 140 150 160 170 180
m237.pep IFALDAAFRVNQPTVLMRFDARHRGLENLRTGFGSGTSDAQSVSERMQVSGXGVELCP

```



m237/a237 85.6% identity in 382 aa overlap

	10	20	30	40	50	60
m237.pep	MRDKVGGNVALPAPRIFDFDIGKLRKNFKHILADKLGHTLRIVDKLVILTAEQSAVRAE					
a237	MRDKVGGNVALPAPRIFDFDIGKLRKNFKHILADKLGHTRGIVDKLVILTAEQSAVRAE					
	10	20	30	40	50	60
	70	80	90	100	110	120
m237.pep	AVIIQNMVAVYCHIVTDKPFCAPOGFGGRNNKGADSNRLAFQRPYRVQTCISIDSIDH					
a237	AVIIQNMVAVYCHIVADKPFCTRAQGFCCGNNKGADSNRLALQRLEYRIQTGISIDGVHQ					
	70	80	90	100	110	120
	130	140	150	160	170	180
m237.pep	IFALDAAFGRVNQPTVLMRFDAHRGLENLRTGFGSGTSDAQSVSERMQVSGXGVELCP					
a237	IFAFDAAFGGVNQPTVLIRFNAYHGRMLKNLRTSFGSGAGDAQRVIERMEMPGQGIELCA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m237.pep	LVHIAVKLLFQLGFIPKLIMTRTMPLGVFMPLQLFPMPLRTDGNRGITALPITIDGMFA					
a237	LVHIAVKLLQLQSVIPELIMSCTVIFLGVLMPLQLFFPMPLRTDGNRGITALPIANGMFA					
	190	200	210	220	230	240
	250	260	270	280	290	300
m237.pep	DAFVHQFDRLQRLLPKPLRLLQADLFFNFPHPTAXVIADNLPATPSRAETDTRGFQHNRF					
a237	DAFVHQFDRLQRLLPKPLRLLQTDLFFNFLHTAGVIADNLPATPSRAETDTRGFQHNRF					
	250	260	270	280	290	300
	310	320	330	340	350	360
m237.pep	MSLLRQGQCSAQTTQSAADDTGIGFQTALKFRINSMRINRTEIIRRQIFLKIRANHCVCF					
a237	MSLLRQGQCSAQTTQSAADDTGIGFQTALKFRINSMRINRTEIIRRQIFLKIRANHCVCF					
	310	320	330	340	350	360
	370	380				
m237.pep	IRCIFGRNDTGCRAISSXQKIGX					
a237	IGYIFGRNDTGCRAISSKQKIGX					
	370	380				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 887>:

```

g238.seq
1   atgaatttgc ctattcaaaa attcatgatg ctgttggcag cggcaatatc
51  gatgctgcat atccccatta gtcatgcgaa cggtttggat gcccgtttgc
101 gcgatgatat gcaggcaaaa cactacgaac cgggtggcaa ataccatctg
151 tttggtaatg ctgcgcggcg tggttaaaat cgggttttgc ccgtccaaac
201 atttgatgca actgcggtcg gccccatact gcctattaca cacgaacgga
251 caggatttga aggtgtttat ggctatgaaa cccatttttc aggacacgga
301 cacgaagtac acagtcctgt cgataatcat gattcaaaaa gcacttctga
351 tttcagcggc gccgtagacg gcggttttac cgtttaccaa cttcatcgga
401 cagggtcgga aatacatccc gcagacggat atgacggggc tcaaggcggc
451 ggttatccgg aaccacaagg ggcaagggat atatacagct accatatcaa
501 aggaacttca accaaaaaaa agataaacac tgttcgcaa gccctttttt
551 cagaccgctg gctaaaagaa aatgccggtg ccgcttccgg ttttctcagc
601 cgtgcggatg aagcaggaaa actgatatgg gaaaacgacc ccgataaaaa
651 ttggcgggct aaccgatggt atgatattcg cggcatcgtc caaggtcgcg
701 ttaatccttt ttaacgggt tttcaagggt tagggattgg ggcaattaca
751 gacagtgcgg taagcccggt cacagatata gccgctcagc agactctaca
801 aggtattaat gatttaggaa atttaagtcc ggaagcaca cttgccgccg

```

545

```

151 DYPPPGGARD IYSYYVKGTS TKTKTNIVPO APFSDRWLKE NAGAASGFFS
201 RADEAGKLIW ESDPNKNWWA NRMDDVRGIV QGAVNPFLMG FQGVGIGAIT
251 DSAVSPVTD TAAQQTQGIN DLGKLSPEAQ LAAASLLQDS AFAVKDGINS
301 AKQWADAHFN ITATAQTALS AAEAAGTVWR GKKVELNPTK WDWVKNTGYK
351 KPAARHMOTL DGEMAGGNKP IKSLPNSAAE KRKQNFEFN SNWSSASFDS
401 VHKTLPNAP GILSPDKVKT RYTSLDGKIT IIKDNENNYF RIHDNSRKQY
451 LDSNGNAVKT GNLQKGAKD YLQQQTHIRN LDK*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 238 shows 86.0% identity over a 401 aa overlap with a predicted ORF (ORF 238.ng) from *N. gonorrhoeae*:

m238/g238

m238.pep	10	20	30	40	50	60
	MNLPIQKFMMLFAAAISLQIPISHANGLDARLRDDMQAKHYEPGGKYHLFGNARGSVKK					
g238	10	20	30	40	50	60
	MNLPIQKFMMLLAAAI SMLHIPISHANGLDARLRDDMQAKHYEPGGKYHLFGNARGSVKN					
m238.pep	70	80	90	100	110	120
	RVYAVQTFDATAVSPVLPITHERTGFEGVIGYETHFSGHGHEVHSPFDHHD SKSTSDFSG					
g238	70	80	90	100	110	120
	RVCAVQTFDATAVGPILPITHERTGFEGVIGYETHFSGHGHEVHSPFDNHD SKSTSDFSG					
m238.pep	130	140	150	160	170	180
	GVDGGFTVYQLHRTGSEIHPEDGYDGPQGS DYPPPGGARDIYSYYVKGTS TKTKTNIVPO					
g238	130	140	150	160	170	180
	GVDGGFTVYQLHRTGSEIHPADGYDGPQGGY PEPQGARDIYSYHIKGTSTKTKINTVPQ					
m238.pep	190	200	210	220	230	240
	APFSDRWLKENAGAASGFFSRADEAGKLIWESDPNKNWWANRMDDVRGIVQGAVNPFLMG					
g238	190	200	210	220	230	240
	APFSDRWLKENAGAASGFLSRADEAGKLIWENDPDKNWRANRMDDIRGIVQGAVNPFLTG					
m238.pep	250	260	270	280	290	300
	FQGVGIGAITDSAVSPVTD TAAQQTQGINDLGKLSPEAQ LAAASLLQDS AFAVKDGINS					
g238	250	260	270	280	290	300
	FQGVGIGAITDSAVSPVTD TAAQQTQGINDLGNLSPEAQ LAAASLLQDS AFAVKDGINS					
m238.pep	310	320	330	340	350	360
	AKQWADAHFNITATAQTALSAAEAAGTVWRGKKVELNPTKWDVVKNTGYKKPAARHMOTL					
g238	310	320	330	340	350	360
	ARQWADAHFNITATAQTALAVAEAAGTVWRGKKVELNPTKWDVVKNTGYKKPAARHMOTV					
m238.pep	370	380	390	400	410	420
	DGEMAGGNKPIKSLPNSAAEKRKQNFEFNSNWSSASFDSVHKTLPNAPGILSPDKVKT					
g238	370	380	390	400	410	
	DGEMAGGNRPPKSI - TSEGKANAATYPKLVNQLNEQLNNIAAQDPRLSLAIHEGKKNFP					
m238.pep	430	440	450	460	470	480
	RYTSLDGKITIIKDNENNYFRIHDNSRKQYLD SNGNAVKTGNLQKGAKDY LQQQTHIRN					
g238	420	430	440	450	460	470
	IGTATYEEADRLGKIWVGEGARQTS GGGWLSRDGTRQYRPPEKKSQFATTGIQANFETY					

547

```

a238      |||||
          FQGVGIGAITDSAVSPVTDAAQQTLOGINHLGNLSPEAQLAAATALQDSFAVKDGIN
          250      260      270      280      290      300

          310      320      330      340      350      360
m238.pep  AKQWADAHPNITATAQTALSAEEAAGTVWRGKKVELNPTKWDVWKNTGYKKPAARHMQTL
          |||||
a238      ARQWADAHPNITATAQTALAVAEAAATTVWGGKKVELNPTKWDVWKNTGYKTPAVRTMHTL
          310      320      330      340      350      360

          370      380      390      400      410      419
m238.pep  DGEMAGGNKPIKSLP-NSAAEKRRQNFEKFNSNWSSASFDSVHKTLTPNAPGILSPDKVK
          |||||
a238      DGEMAGGNRPPKSITSNSKADASTQ

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 893>:

```

g239.seq
1  atgttcacc ataaaggtat tgcccgaaac cggcggatgg aggttttgtt
51  tttctgccgc cgccctgatc gcttcgtgat tcgcccacac gcgctgttgc
101 agcctcattt gcgcataatc ctgctccaag gcgatttcct gttttttcgc
151 cttgtccaaa gctgtgaagt tgagcctgta ctgggtttgc tgcatacaca
201 cggaaaaagc ggaacgcac accgcaagca gcagaaagaa attcgatttg
251 ttcattgccg ttcagacgtt tttctctgtt attattccgg tatcggaaccg
301 gcagtcgcgt ccgcccacac caaaactgcg ctctctgccc tcgggttggc
351 ggcaatttcc gttcaccggt gctttaatgc cctgccacac attttcaggg
401 gcggatcggt caaatccgct tctctgaccg ccgcccagct cggcaggggc
451 tcgtgttgcg aatatttttt gacaaactgc ttcacaatgc ggtcttccaa
501 cgaatggaaa gcaatgaccg ccaaacgccc gccctctttc agacggcaca
551 tgacctgcgg caataccgcc cctacttctt caagctcgcg gtttaataaag
601 atgcggattg cctggaaggt gcgcgtcgca ggatccctgcc ccgctcgcg
651 agtacggacg tttgtgcca caatctgcgc cagcttgcgg gttgtatcga
701 ttggactttc cgcccgttgc gcgacaatgg cgcgcacaat ctggcggcta
751 aaccgctctt caccataa

```

This corresponds to the amino acid sequence <SEQ ID 894; ORF 239.ng>:

```

g239.pep
1  MFHHKGIARN RRMEVLFFCR RPDREIVIRQT RLLQPHLRII LLQGDFLFFR
51  LVQSCEVEPV LVLLHHNGKS GNAHRKQKE IRFVHCRSDV FLCYYSGIGP
101 AVRSATRKTA LLALGLAAIS ASPGFNALPT IFRGGSGKSA SLTAAQLGRG
151 SCCEYFLTNC FTMRSSNEWK AMTAKRPPSF RRHMTGNTA PTSSSSRLIK
201 MRIAWKVRVA GSCPRSRVRT FCATICASLR VVSIGLSARC ATMARTIWRL
251 NRSSP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 895>:

```

m239.seq
1  ATGCTCCACC ATAAAGGTmy kGCCCCGAAAC CGGCKGATGG AGGTTTGTGT
51  TTTCTGCCGC CGCCCTGATC GCTTCGTGGT TCGCCAAACG CGCCTGTGTC
101 AGCCTCATTT GCGCATAATC CTGCTCCAAG GCGATTTCCT GTTTTTCGCG
151 CTTATCCAAA GCTGTGAAAT TGAGCCTGTA CTGGTTTTCG TGCATCACA
201 CGGAAAAAGC GGAAACGCAC ACCGCAAGCA GCAGAAGGAA ATTCAATTG
251 TTCATTGCCA TTCAGACGTT TTTCTCTGTG ATTGTTCCGG TATCGGACCG
301 GCAGTCCGCT CCGCCACACG CAAAACCGCA CTTCTCGCCC TCGGATTGGC
351 GGCAATTTCG GCCTCACCCG GCTTTAATGC CCTGCCACG ATTTTCAGGG
401 GCAGCTCGGG CAAATCCGCT TCCCTGaCCG CCGCCACGCG CGGCAGGGC
451 GCGTGTGCGG AATATTTTTT GACAACTGCG TTCACAATGC GATCTTCCAA
501 CGAATGGAAA GCAATGACCG CCAAACGTCC GCCCTCTTTC AGACGACACA
551 TGACCTGCGG CAATACTGCC CTAATCTCTT CAAGCTCGCG GTTAATAAAG
601 ATGCGGACCG CCTGGAAGGT GCGCGTCGCA GGATCCTGCC CCCGCTCGCG
651 AGTACGGACG TTTTGTGCCA CGATCTGCGC CAGCTTGCGG GTTGTATCCA
701 TTGGACTTTC CGCCCGTTGC GCAACAATGG CGCGCGCAAT CCGCGGGCTa
751 AACCCTCTT cACCATAA

```

This corresponds to the amino acid sequence <SEQ ID 896; ORF 239>:

This corresponds to the amino acid sequence <SEQ ID 898; ORF 239.a>:

```
a239.pep
1  MLHHKGIARN RRMEVLFFCR RPDRFVVRQT RLLQPHLRII LLQGDFLFFR
51  LIQSCEVEPV LVLLHHNGKS GNAHRKQOKE IQFVHCHSDV FLCDCSGIGP
101 AVRSATRKT LALGLAAIS ASPGFNALPA IFRGGSGKSA SLTAAQRGRG
151 ACCEYFLTNC FTMRSSNEWK AMTAKRPPSF RRHMTGNTA PTSSSSRLIK
201 MRIAWKVRVA GSCPRSRVRT FCATICASLR VVSIGLSARC ATMARAIWRL
251 NRSSP*
```

m239/a239 97.3% identity in 255 aa overlap

	10	20	30	40	50	60
m239.pep	MLHHKGXARNRXMEVLFFCRRPDRFVVRQTRLLQPHLRIILLQGDFLFFR	LIQSCEIEPV				
a239	MLHHKGIARNRRMEVLFFCRRPDRFVVRQTRLLQPHLRIILLQGDFLFFR	LIQSCEVEPV				
	10	20	30	40	50	60
	70	80	90	100	110	120
m239.pep	LVLLHHNGKSGNAHRKQOKEIQFVHCHSDVFLCDCSGIGPAVRSATRKTALLALGLAAIS					
a239	LVLLHHNGKSGNAHRKQOKEIQFVHCHSDVFLCDCSGIGPAVRSATRKTALLALGLAAIS					
	70	80	90	100	110	120
	130	140	150	160	170	180
m239.pep	ASPGFNALPTIFRGSSGKSASLTAAQRGRGACCEYFLTNCFTMRSSNEWKAMTAKRPPSF					
a239	ASPGFNALPAIFRGSGKSASLTAAQRGRGACCEYFLTNCFTMRSSNEWKAMTAKRPPSF					
	130	140	150	160	170	180
	190	200	210	220	230	240
m239.pep	RRHMTGNTAPTSSSSRLIKMRTAWKVRVAGSCPRSRVRTFCATICASLRVVSIGLSARC					
a239	RRHMTGNTAPTSSSSRLIKMRIAWKVRVAGSCPRSRVRTFCATICASLRVVSIGLSARC					
	190	200	210	220	230	240
	250					
m239.pep	ATMARAIIRRLNRSSPX					
a239	ATMARAIWRLNRSSPX					
	250					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 899>:

```
g240.seq
1  atgatagaag tcatacatTT cttcggcgcc gaaacgcgca gacagtttgc
51  ttgtgccgac gttggacgat ttctgcataa tgccgcgcac atccaaagag
101 gggtaaakat gggatatcat gcgcacggga gacggtcgca ttttataagg
151 ctgcgtatcc agccgttcgt tcaaatcggt ttgcccgcga tccaatgcct
201 tcgcaatcac gaacggtttg attgccgaac cagggtcgat catatcggtt
251 acggcacggt tgcgcgcgtg ttcgctgtct gcccggcgcg gtctgttggg
301 atcgtaggcg ggcgtattgg ccaaggcgag gatttcccc gtgcgggcat
351 ccaaaaccac caccgttccg gcttttgctt gatggtattc gaccgccttg
401 ttcaactctt cataggccaa ggtctgaatc ctctgatcga gggaaaggat
451 gatgtctttg ccgttttgcg gtgctttatt gcgcggggag tccaagctgt
501 ccacaatatt gccctgccgg tcccgcaaaa caacttcgcg gccgtcttcg
551 ccatacaggc tgtcttcaag cgaaagtcc aaaccttcc gacctttgcc
601 gtcaatatcg gtaaatccga tgacgtgtgc aaacagggtt cccatcgggt
651 aatggcggtt taa
```

This corresponds to the amino acid sequence <SEQ ID 900; ORF 240.ng>:

```
g240.pep
1  MIEVIHFFGA ETRRQFACAD VGRFLHNAAH IQRGVNMGII AHGRRSDFIR
51  LRIQPFVQIG FARIQCLRNH ERFDCRTRFD HIGYGTVAPL FAVCPAGSVG
101 IVGGRIGQGE DFPRAGIQNH HRSGFCLMVF DRLVQLFIGQ GLNPLIEGKD
151 DVFAVLRFCF ARGVQAVHNI ALPVPQNNFR AVFAIQAVFK RKFQTFLTFA
```

551

```

251 GCACGGTTGC GCCGCTGTTC GCTGTCTGCC CGGCCGGGCC TGTGGGGATC
301 GTAGGCGGGC GTATTGGCCA AGGCGAGGAT TTCCCCCGTG CGGGCATCCA
351 AAACCACCAC CGTCCCGCT TTTGCCTGAT GGTATTCGAC CGCCTTGTTT
401 AACTCTTCAT AGGCCAAGGT CTGAATCCTC TGATCGAGGG AAAGGATGAT
451 GTCTTTGCCG TTTTTCGGGG CTTTATTGCG CGGGGAGTCC AAGCTGTCCA
501 CAATATTGCC CTGCCGGTCC CGCAAAACGA CTTCGCGGCC GTCTTCGCCA
551 TGCAGGCTGT CTTCAAGCGA AAGTTCCAAA CCTTCCTGAC CTTTGCCGTC
601 AATATCGGTA AATCCGATGA CGTGTGCAAA CAGGTTGCCC ATCGGGTAAT
651 GCGGTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 904; ORF 240.a>:

```

a240.pep
  1 MIEVIHFFGT ETRRQFACAD VGRFLHDAAH IQRGVNMGIA HGRRSDFIRL
  51 RIQPFVQIGF ARIQCLRNHK RFD CRTGFDH IGYGTVAPLF AVC PAGPVGI
101 VGGRIQGED FPRAGIQNH RSGFCLMVFD RLVQLFIGQG LNPLIEGKDD
151 VFAVFRGFIA RGVQAVHNIA LPVPQNDFRA VFAMQAVFKR KFTFLTFAV
201 NIGKSDDVCK QVAHRVMAF*

```

m240/a240 99.1% identity in 219 aa overlap

	10	20	30	40	50	60
m240.pep	MIEVIHFFGTETRRQFACADVGRFLHDAAH	IQRGVNMGIAHGRRSDFIRL	RIQPFVQIGF			
a240	MIEVIHFFGTETRRQFACADVGRFLHDAAH	IQRGVNMGIAHGRRSDFIRL	RIQPFVQIGF			
	10	20	30	40	50	60
	70	80	90	100	110	120
m240.pep	ARIQCLRNHKRFD CRTGFDHIGYGTVAPLFAVCPAGPVGIVGGRIQGEDF	PRAGIQXHH				
a240	ARIQCLRNHKRFD CRTGFDHIGYGTVAPLFAVCPAGPVGIVGGRIQGEDF	PRAGIQNH				
	70	80	90	100	110	120
	130	140	150	160	170	180
m240.pep	RSGFCLMVFDRLVQLFIGQLNPLIEGKDDVFAVFRGFXARGVQAVHNIALPVPQNDFRA					
a240	RSGFCLMVFDRLVQLFIGQLNPLIEGKDDVFAVFRGFIARGVQAVHNIALPVPQNDFRA					
	130	140	150	160	170	180
	190	200	210	220		
m240.pep	VFAMQAVFKRKFTFLTFAVNIGKSDDVCKQVAHRVMAFX					
a240	VFAMQAVFKRKFTFLTFAVNIGKSDDVCKQVAHRVMAFX					
	190	200	210	220		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 905>:

```

g241.seq
  1 ATGATAGAAG TCATACATT CTTCGGCACC GAAACGCGCA GACAGTTTGC
  51 TTGTGCCGAC GTTGGACGAT TTCTGCATGA TGCCGCGCAC ATCCAAAGAG
101 GGGTAAACAT GGGTATCGCG CACGGGAGAC GGTCCGATTT TATAAGGCTG
151 CGTATTCAGC CGTTCGTTCA AATCGGTTTT GCCCGCATCC AATGCCTTCG
201 CAATCAGAAA CGGTTTGATT GCCGAACCGG GTTCGATCAT ATCGGTTACG
251 GCACGGTTGC GCCGCTGTTC GCTGTCTGCC CGGCCGGGCC TGTGGGGATC
301 GTAGGCGGGC GTATTGGCCA AGGCGAGGAT TTCCCCCGTG CGGGCATCCA
351 AACCACCAC CGTCCCGCT TTTGCCTGAT GGTATTCGAC CGCCTTGTTT
401 AACTCTTCAT AGGCCAAGGT CTGAATCCTC TGATCGAGGG AAAGGATGAT
451 GTCTTTGCCG TTTTTCGGGG CTTTAKTGCG CGGGGAGTCC AAGCTGTCCA
501 CAATATTGCC CTGCCGGTCC CGCAAAACGA CTTCGCGGCC GTCTTCGCCA
551 TGCAAGCTGT CTTCAAGCGA AAGTTCCAAA CCTTCCTGAC CTTTGCCGTC
601 AATATCGGTA AATCCGATGA CGTGTGCAAA CAGGTTGCCC ATCGGGTAAT
651 GCGGTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 906; ORF 241.ng>:

```

g241.pep
  1 MPTRPTRAAN PPTPTWLQT AYCPRPPYRP PSVQHTPHE PASSTCAAKS

```